

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACCTC
AGTGACTGAGGCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAAGCTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCTCGCCGCTGTG
CCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCCGTACCTGACCTCCATCCCAG
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAAGCTGGAGTTACTCATGCTGCACAGT
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAAGTCTTGCAGGTC
TAAAAAATGAGCTATAACAAAGTCCAAATCATTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTCTCTATGTACAA
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAAGTGAACGCACTGTCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCCTCAGGGGGCCGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACTCCAGTGTTCCTTAAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAAACATAGAGACTTAATACATCGGCGGCGTGGGGATTCCAGC
CTCCGGCGATTACAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAACCTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCCAGATGAAGAATTCATGGTTC
TGAAAACTAAGGCTTCTGGTGTCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCTGATTTCAAATTATTTAGTGTAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTACACACCACTGCCTCGTCTTTATTTACATTCTTAGAAACAA
CAATACAGGTAACCTCCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGAGAGTTAAAAACCCACATAGAACCCAGTTCTCCGACGGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCTTCCGCAGAGGGGGCTCACAG
TGGCTACTGCAGCACTGTGAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAACCTA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAAACTGGGAAAGATTAGTGGAACAACACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACCTCCCTTG
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAGTCCTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCCTTGGCTTTTCAAGTCATTCCCCACGACAAGATGATGGTGGAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCACTCCTCAGC
AGTGGGGCGCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC
TGGTCAAAACCCAACCTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAATAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACCTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCACT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT
ATCTGCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA
TTTTGGTTGCTGCCTTCCAACAATGTCATTTTCTCCAATGACAGGTTACATTTC
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATG
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAACTGGACATT
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAGCC
ACAGCCATTCGGCACTCCAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT
TCCCAGGTCACGTGGATTATGCCAGGCAATATTTTCCTCCCAGCTCCATACTTTGGA
AGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAGATGAGGAACATCCGGCTTTC
TGACTCTGCGGACTTCACCTGTGTGGTTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT
AGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTGAGAAACCCATTCA
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACTGCTCTGTGGAT
GGGAACCCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC
AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA
GCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
CATCGAGAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC
AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTG
GCCAGGCAGTTATTAGCGTGTCAAGTGTGGTTGTGGCCTACCCTCCCCGAATCATAA
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT
GCTCTCAAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG
GTACGCTGGTCATTCAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
CTCAGAACCTACTTGGGACTGATTACGCAACAACCTTACATCCAGGTACTCTGACAGG
AAGGGGGAGACTAAAATTCAACAGAAGTCCACATCCACAGGGTTTATTTTTTGGAA
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT
ATTAAATTTACAATGGACATGCGATGAGACTTGTAATGAAAGCATTGTGAACTGA
AACCGAGTCTCTGTGGATCTCAAAGCAAACCTTAACTTAAGGCACTTTGATTTTGC
CAACAAATAATAACAAACATTAAGAGAAAAAAATGATCCACTACGAAATAACAAAC
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC
TGCCTCGTGTCTGTTTCCTACCAATGTCACAAACATCGCACACAGGGTGAATGGAGT
CAACGGGAAAGATTAAGTTTGGCGTCTGTGTAAATCTCAATGTACAAATATTCTGTC
NCTGGTTTATAAACATTTTGATAAAACCGAAAAA
AAAA (SEQ ID NO: 1)

Figure 2

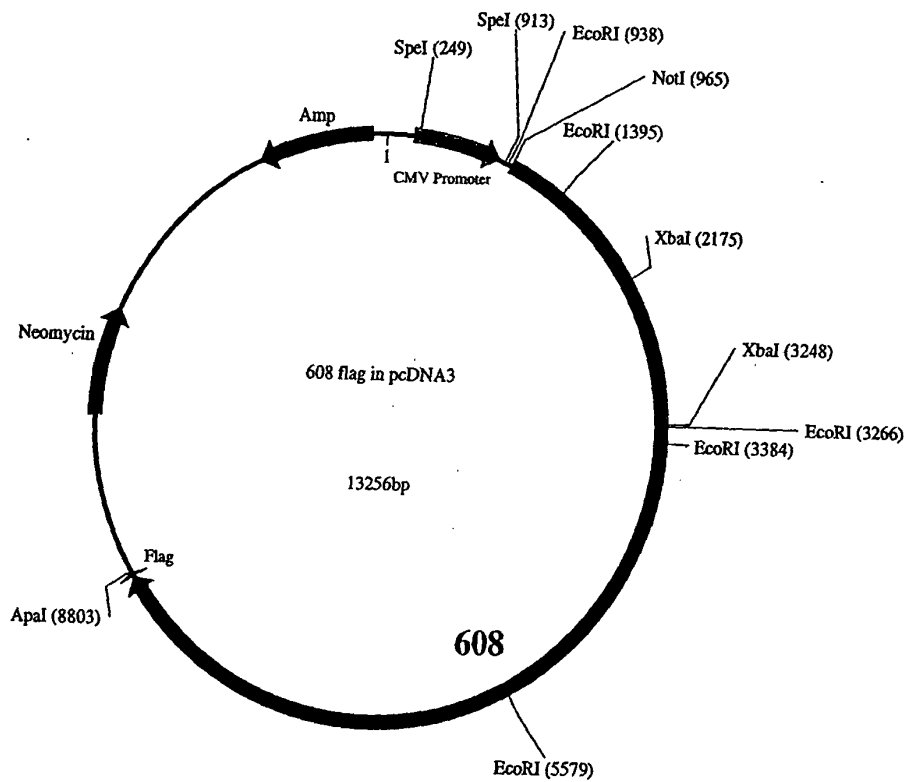


Figure 3

MQVRGREVSGLLISLTA VCLVVT PGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNGS
ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
ADPFWFQQEKIVLQLNRTATTLSLTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNNPK
LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
DSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVGTGETLDLPLCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRLQVTPKDQGHYQCVAAANPSGADFSSFKV
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSASALTGSEAGKQVSGVHRK
NKHRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
KPVPLAVPLVELTDEEKDASGMIPDDEEFMVLKTKASGVPGRSPTADSGPVNHGFMSTI
ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE
IRDSAQAGRASSQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQIPITGV
SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
KNPHRTPVLRHRHRHRTVRPAIKGPANKNVSQVPATEYPMCHTCPSAEGTLVATAALS
VPSSSHSALPKTNNVGVI AEESTTVVKKPLLLFKDKQNV DIEITTTTKYSGGESNHVIPTE
ASMTSAPTSVSLGKSPVDNSGHL SMPGTIQTGKDSVETTP LPSPLSTPSIPTSTKFSKRKTP
LHQIFVNNQKKEGMLKNPYQFGLQKNPA AKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
TMAATQNGKTEVVSGARSL SAGKKQPFTNSSPVL PSTISKRSNTLNFLSTETPTVTSPTAT
ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALTTADTPLAFSHSP
RQDDGGNVSAVAYHSTTSLLAITELFEKYTQTLGNTTALETTL LSKSQESTTVKRASDTP
PPLLSSGAPPVPTSPPPFTKGVVTD SKVTS AFQMTSNRVVTIYESSRHNTDLQQPSAEAS
PNPEIITGTTDSPSNLFPSTSVPALR VDKPQNSKWKPSWP EHKYQLKSYSETIEKGK RPA
VSMSPHLSLPEASTHASHWNTQKHA EKS VFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
GKAASFTVPANSDVFLPCEAVGDPLPIIHWTRVSSGX EISQGTQKSRFHVLPNGTLSIQRV
SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
RPTVSWILANQTVVSETAKGSRKVWVTPDGT LIYNLSLYDRGFYKCVASNPSGQDSLL
VKIQVITAPPVIEQKRQAIVGVLGGS LKLPCTAKGTPQPSVHWVLYDGT ELKPLQLTHS
RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVVILTVEEGETIPRIETASQKWTEVN
LGEKLLLNC SATGDPKPRIIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD
YLCVARNKMGGDDLVMHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFNNVGMAEEGDYICSAQNTLGK
DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
NDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
VIKATAIRHSSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

1000 900 800 700 600 500 400 300 200 100 0

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLILYKATRKNKSGKYRCAARNKVGYYIEKLIL
LEIGQKPVILTYEPMVKSVSGEPLSLHCVSDGIPKPNVKWTTTPGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGNVICRAQNSVGQAVISVSVMVVA YPPRIINYLPNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240

Figure 4

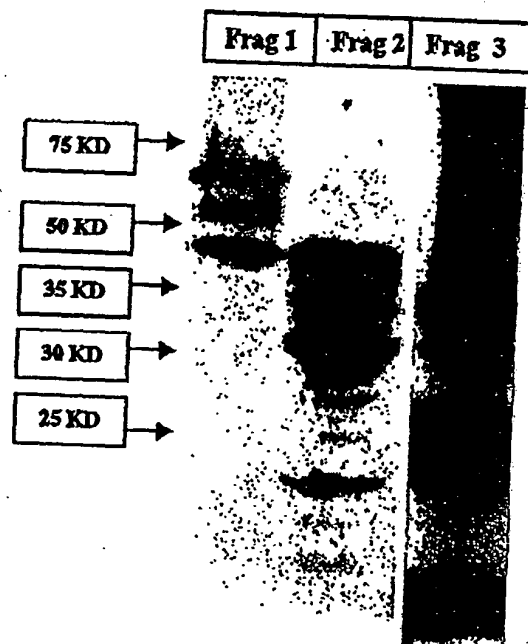


Figure 7

Nuc 1

↓
TTTGGAAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTGCATTTA
TCCAATGGAATACTACTCAACTTATTAACCAACGACTTTCATAAAATTTTAGGCAAAAT
GNATGGTCTGNAGGATCTTGAGTGAGGTAACCAATCACAAAAGAACACTCATGGTATGC
ACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTIAAAAGGGAAGAAGACACATAGC
TTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGTTGGCTTATGAGTGAATCT
TGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAACGTAGGAAGAAGTAGTTAATCA

↓ Nuc 390

AGGTGACAAAAGTATCCTGACACATTGGGAAAAGACCACAGTCCAGGAACTGAGTCTTAA
GGATTTCATATTAACCTCAGTTCCCATGTGCAGCTCTGAGACTTTGGCAGATCAGACACTT
AACTTCACCAGCTTCCTACACAGAGCAGTTACTATCCTTGCACCTTCACAGATGGAGTGTGA
CCATTAAAGCTGCACGTGAAACATGAGTCTGACTTGTAAATAATCTTAAAAATACAAATTGTGT
TGTAAGATATGTGACCAAAAGAGCATGGTCACTCTATTAACCTTTGATGTTCTATGGACTCT
TAATTTTATGGTATAGAAATGTCAACAAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAG
GATGGCCTTTTCAGTTTGTAAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCA
ATTAAACCCOCCAAGTAGGCATAATNTAATGCTTACTTCATCAGAATATGTCTAATTGAC
TCTTCTAAAAAGACTTTGGTATGCTATAGGATCTAAATGTAAATGTGATTACTGACATAAT
AAATAGGAQAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTCTAATAGGTCTT
ATAGGTTATATAAGAGGTGGGAAAGGAATATTGAAACATCTAGAAGTAAAATGATCCCTG
AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTGTGGTTTGGCT
GGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATACAGAGCTAT
GACGTTAGTCTGGAAGGAGCTTTGGAAGAAATGACAAGCTGTAGCTGCCAGAACATACTA
GATGCCATATTTCCAAGGCAAGTGTCCACATGGGGACATCTTAAGAATATGGTTGTCTCTG
CAGTGCTAAGGACCTTGTGTGCTGCCACACAGGTCTCCAGGTTAGTGCTAACTCTGACTGC
TTGACTCTTTAATTTCTACCTTGATCATTAATGACTAGAAATCACTTGGTGTATTAGCAACTG
GATATGGAATATTACTAATTTGTACCCAAGCCAGGCCACCTCAGCTTTGGCAGCTCCATTG
ATTCTGTGGAGCCCACTGACGTTGGGTTTGAATCAACTGTACTGTTCTACTTACAAGAGCG
ATTACCTGAGATGAGTCAATTTTCTTCACAAGTCTTTTGAAGAGTCAATPAGACATATT
CTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATGTGTTCATGTATGCTCATGGA
TGCTATTATAATGTGGAAATAAACTGACTTTAAAAAAGGCTTATGATACITGTTCACA
QAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTA
AAAGCTTTAGAAATTTTGTAAACAAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTGTA
TGGTGGGCTTGAGAGATGGAGCTTAAGAACTTACTTGTCTCCAGGAGCACATCTTCAGAAC
ATCTGCTCAAAACATTTATCCCAAATGCTCATCAAAAGGCTCACTCACATGTGCTTCAACC
ACAGGGATTAAACAGTCAATTTAGTCAATTTCTCAAAAGGCTGGAAGGCTGTAGAGGAA
CAGGATGTATCAGGATAACATCCAACTTACAAAAGGATGTCTATAACCTCAACCAACA
AACAACAACGACAACAACCCATAAAATTTATCAAGGCAATGAACATAAGCCATATGCA
GAAAAAGTATTATATGTCTCATTTGTGGGGTGTTTTCTTAATAGTCAAAATATGCAGAAT
ATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTTGTAGATTAGAGGACA
CAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTTAATGTATAATATGAGGACTGTA
TTAATAATTCTATTAAAGATACACAGCAACGAGTGTATCTTACTAACACACACACTTAC
ATAGAGAGAATAAAGTGATAGATACGTTTGTTTTATCTTCATGTAGCTGATAATTTCATAT
TGTACACCTCAAAACATAGATAACCAACAAGAGGAAAGAGGATAGGTGCCTCTCCAGGGC
GGAAGAGTACATTGGAAGTCAAGACACCATTTGTGTAGATGTACCACATGGAGGAGTAGA
GAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCTTATAGGTGGAACAACATTATGAGCT
AACCAGTACCCCGGAGCTCTGACTCTAGCTGCATATATATCAAAAGATGGGCTAATCCG
CCATCACTGGAAGAGAGGCGCATTTGACTTTGCAAACTTTATATGCCCGAGTACAGGGGA
ATACCAGGGGCCAAAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGG
GGACTTTTGGTATAGCATTGGAATGTAAATGAGTTAAATAOCTAATAAAAAATGGAAAA

↑ Nuc 2317

AAAAAAAAAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTT
GCAAAACAGAAAGGGGAGTCACAAACCCAGGACAAACACAGTGATTGAAGCGTCTTTGAA
TGTTATTGCTGTTGTTGTTACCAACCATATTAGCATATATTCAATTGTGAAAACCTTACGGGT
CTATGACATGTTTTTTATTCAAGTATATCACATGCTGTGACGATATTGGCACCCTACCA

[illegible]

GCTGGGGTTTCAGGGCGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAAAGGAAAGT
ACTGGAGGGGAGTTGGGACAAAAGCAGCGAOC AAGGGAA CATCGCTTCAGTGACTGAAGCCA
GGCAAAAGGAGCGCGGAAGGATTATATGTAGCTCTGGGAOGCTTCATAAACACTGATGAOGTGT
TTGTGCAAAAGCAAGCAATTTGAGGAGAAAACGCTGGGAOGCTCGCAAGAAAGGAGTGATCGATT
AGTACTTGTAAAGTTTAGGTGAGTTTGAGAACTAACTAACTATACTATTGAGGGAGAAAGCAAG
AGCATTTCAGCAGCAGCAGCAGCAGCAGCAATCAGATAAAAGGAAAGCTTTGGTTAGTTTGGAA
ATGTATGATACCATTAATAAATAACAGAAAGOGCTCCAGTTCTCTGAAGAGTGAGTCCCCAGCTA
GTGAAGACTAAAGCTACTAAGGCTTTTGCTCCCGTTGGAAGCAAAGAAAGTTTCTTCAATCAGG
TGAAGGCTCTCTCAGAAAGATTCTGTCTGTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGAC
AGAAGAGCTCAGGTATTGCCAACTCTTT

TGTTAAATACAGTTTGAGGCTTAAGTGATACGGGAACTCATGTGGTATTCAHTAACGGCTCT
CTTCTCTTATAACTAAGCTTAAGGTGCATATAGTCTCTTCTGTTTCCAGCTACCTTGTAC
CACTCTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATACCCAGAGAGTATT
CAAAAATTGTCAGTGATGTAACAGGTGACAGCTGTAGGCATAGAAGTAATCATTAGTAAATC
TTAATTGGGTTTAAACATTCATTCATAAGCTCCAGGTGGGAGGGATCACTGAGGCTCTCC
CACGTGGGGTTTAAAGCATATTTTCTAAACAAGAGAAGCAGAAATTTCTCTCTGGCCATGCTCC
CCATCACTGTGTCACTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAGAGCAGACAGTGT
TGCCCTGCAAAGTCAGAGACTCAGGCTCCAGCTGGTCACTTACTGTCTCCCGGTCAAT
AGTTGGCTCTGAAAAGGCCATGTGTCTTATTGGCAAAGGACTTGCAGACATGCTAGAAA
GAAATTTGACCTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGGAAAGGTTTAAAG
CCAAATAAATAAAACATATTAATAATACATGTTTACAAAAATTGATCATATAAGAA
GTAATTCATAAATGCAATGTGAAAAATATATAATTTTATCTATTCTGGTGC
TTTCTAAATTCAGCATGTACTATTTTATATTTATAAAAAATTTTAAAAATGTATATAAA
AGTGTAAAGGCTCTGTGTCAAACAAGAGAGTTAAATTTACAAACTTTAATTGTCCGAT
AACATTATTATGATCTCTAATGACAGGGATCTGCTTTTCATTGGGAAATGAGAAGCTATG
AAGATATGTTTACAATAATAAGCCATTAGTGATAAAGTCCAATGGGAAGCTAGCACAC
ACTGGTTTATAAAGGAACAGTTTCTCTGATCTATGCAAGTTTACACTAGGGAATAAG
AGTTCCTCTTCTCCAGATTTCATGCAATTTGTTGATCATTTTCTCTCTGATGATGAG
CATTATAAGCTGGAATAAGATAGGATCTCAAAAGGAATGTCAATTTGGATGCCCTGAACAA
CTTTCAGGCTCTTCTTTCAGTTTCACTAGTCTATTTCATTTATTGGATAATTGGGGGATGGTGT
TAATTTTGTGAGTCTTATGGAATTCAAAAAAACAAAAACAAACAAACAAACAAACAA
ACCTCTGAAAGCTAGAACTACCAATCCATTACTGGGTATGTAAACAAAGAGAAATCTGCACA
GAATTTATTGCTACATTGTTTATTTCACCAAGCAGCCAAAGAATGTGGAACCAACTTACGTA
GCCGTCAAAGATGAAOAGATAAAGAAATGTGGAAATGTGTACAACAGAGTCCCATGTG
GGCTAAATCAAACAAGACAGACTCAAAGGAAACACCGTGTAGCTTCTCTGACAAATG
GAAGCTAGATTTACACTTGTACGTGGCATGTGTGTTTGAATTTTATTAGTTATACACT
ATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAACAAGCTAGCTGGG
TGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAACATAGGTG
CCAGGATGAAATGCATTAATTGTATGCTAACTAAACCACAGACAGGAGGCCACAGTTCA
AACCAGGGTGAAATCCAGACAGAGAGAAGGGGAAGTAGACACAAGTTTGGCCACTAAC
CAAGAAGCCATTTCAGTTGCTGCTGCTGGGAGGGGCGTCCAGTTTCTCCAGTCTGAC
ACTGTGTATAACAAACAGTTGACAAATACAAAGTTGGCATGATGATGGTTTTGTGCTATT
TTTCATTTTCTTACTGTTTGTGTGTGTGGTGGTGTGTGTGTGTGTGGCTGTGGTTTC
ATTTGTTCTTTTGAGAGAGAGAAGGAAACATGAAATTGGGTGGGTAGGAAGCTGGAAGCG
ATCTGGAAGAGTTTGGGGAAGAGAAAATTTGATGAGCATATTTAAACAAACAAACA
AAACAAACAAAGGTTCAATTTGCCACAAAAGGTTGTAATTAATTAACAGTTACGACT
CTTAAAGAAAATATTCCCAATTATTCCACAGAGTTGCTATGATGCTGTGCCTAGGACTTGG

▼ Exon 2 (Nuc 6559)

CTTGAACCTGGGOCCTATAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAAGAGAGGCAGGGAA
GTCAGCTGCTTGCTGATCTCCCTCACTGOCATCTGCTGGTGGTCAOCCCTGGGAGCAGGGTC
TGTCCTCGCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTCGGGACCTGACC
TCCATCCAGACGGGCATCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

↑ Nuc 6755

GGACCTTGCTGATCTCCTTCTCAGAGAGGGAGCCACTGATTTTCTGTAATTTGCCCCC
AAACAACCTGTGATTACTTTTAATAGTTTCTTCTAAAAATGGGTTTCATACAAACCTTATATTG
TGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACCTGAAGCCCCCTTTAGTT
GTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTTAAATGAAAAAGAT
ACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAATTGTCTATTGAATATTTTAA
AAAAGCATTTCCTTTTGCAATTTATAAATACCCATTACAAAATGGCTTACTTAAATACTTG
CCTTACTAAATCTGACAAATTATGGTGATATTTTGAAGGTTTATGAAAAATTTGTTTATGTGT
ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTGCCATTAGTGAGCATGTACAGT
ATGCCAAACACTATTGTTTACGTTTGGAGGAAGTAATGGGGGTGGGGGAGCAACAAGGGT
TATAACCGTATACCCAGTGCCCTGGGAAGCGATTGCAAAACAGTAAAGACTGACATTGTGTT
CTCCCTATGAGGGAGGGGCTTGGGCTGAGCATTGCAATGAGCATTGCTCATTGTGCT
GGCAGGTTTATGATAACTTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTTCAACT
GAGAACATGCCGGAAGAAGATCAGATTATAGGCAGGCTGTGGGGCATTTCCTTAATTAG
TGATTTCATGGGGCAGGGCCAGTCCATTGTTCTGTGGTACCATTTCAGGCACATTTAAAA
AAAAAAGGAGGCTGAGCAAGTGTCAAGGAGCAAGTCACTGAGCAGCAGCCCTAATGA
TCCTGTCATCAGCTCCTGCCCTCAGGTTCTACCTATTTGAGTTCTGTCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCAACTTGCTG
TTGGTCATGATGTTTCATCAGATGATAATAGTCTCATGAAGATGCTGGTGTTTATAACA
CCTTTGGACTAAATTCGTTATCTATAGCTGAGGAAAATGGAGCATAGAAAGTCTCCAGA
CTACACCAGAGTGTAATCTGGGCTGAGCTTAGAATCACACCCACGTGCCATCCACTGCC
GGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCTGTTTGTTCATTTTGACA

Exon 3 (Nuc 8089) ▼

TGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATGATTTTTTCCAGGTATAACAGC
CTCACTAGATTGACAGAAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATGTGCACA
GCAATGGCATTACAGAGTCAGTGACAAGACCTTCTGGGCTTGCACTCTTGACG

Nuc 8218 ▲

GTGAGATAGGTAGAGGGTGATGGAGGCTGAGAAGAGAGGTTGCAACTGTGGGTATATCCC
AAAAGCTGCTGATTCCCGTGGGAGACATTTCTATAAGCATTCTATAAACTAGAGGCAGATA
TCAAGGAAGGATTTCATTTGTAATGCAATTTTATGAGAAAATTGAAATTAAGAAAATG
CTGGGGAATAATGCTTACACAATGGGAGGACCTAATTTAGGATCTCCAATAGCCACATAA
AAAGCACAGCATGGGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACTGTTTCAAGT
CCCAGAGACTCATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCACTGACAAA
ACTTGACTCAGAAAATAATGTGGAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACT
CATGCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAGGAAGGAAGGAAGGAAGG
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
GGAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
TTTTATCCATAAAAGGTCATTTCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT
AAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAAATTTTGAGTTTCTCTAC
TTTTATCTCTACCATCTTCAAACCTGAACTACAATTATCCCAAAATGAAGAAAATGC
TGTAAGAGTTTTACACACCGAAGTGGGAACTTAAGGATTAGACAAGTCTAACAAATGAG
AATGGGGAGAAACAAAAGAGACTGCACAGGGAGCCCTTTCTCTGCTTATAATCTTGACAC
TTGAGAAGCTAATTGACGCTGCATGACTACTCAACTCTTTAAGCAAAACATGCTGTTGTT
ATGAAAAGCACAATAAAGTACATATGTCCATAATATTCATCAAAATTTGCATGCAGCAC
ATAATAGCAATCAAAGCAATAACACCCACTGTTACAGAGACTTTAAACATGAAACTGGA
ACTATGTCTAGTGTGTTGACTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTT
ATTTAGGTCATCAGACAGCATTGGGAACATGTATCTTCAGGAGGAATCATCTATCTCT
GCATGAAATTTCTCCACTATGTTTATTCTCTTAGCCAGGTTTTCTCTGATGGAGAAACATT
GGGTTTGAGGTTTACTCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGG
CTTTTATTTATGAGGGATGTTGGTATTCAGAAAATTTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAATATGTAAGAAAGTTATGTTTTATTAGTATTTTCATGTTCTTTCTT
TTTTAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGAAAATAATC
AGAGAGGGATTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTTTTGCTCTAATT
CAGAATCAATTAATTCATATTACTATAAAGACAGCACGCCAGATGTGTGCCAGCTGAG
GAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACAGAAAAGGAAGTGAAGGGTTGA

TGTGGGCTGCAACATCTTGAAAAACATTGGGCTACATGATGGAAGCCAGGCACAAAAAGCC
ACATATTGCATGGTTATGTTTATATGAAATGTTTAAAAATACATGGATTCTTAGCAAACAGA
GTAAGATGTTACTTAGGGTCAGGAAAAAGATTAAAAAATACTATTGATGTGGAATG
ATCTTAATTTGGGGAAAAAGACAATTCCTAAGACGAAATAGTTGAGGTAGATATAGTTAT
ATCCCTGTGGATATTGTAATAAAACAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGG
GCAGGAGGAAAGTGAATGAGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTGCTG
GGTGGAAATCTAGATATGTTAATATATTGACATAAAGGAAGGAATGTTTAGGGAAGGATC
AAAACCAACAGGAGTGAAGGGAGACAATAGGAACCAATGAGAGGCAAAAGTTCATGGTCAA
TGTGTGTGGAGACACCATAATAAAACTCCTTTTTTGTGCTAACTAAAAACCACTAAAATC
TAAAAACAAAACATTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGGG
GAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTAACTGCTTTTACGTTT
TTTATTATTAAATTACTCTGTCTGTCATTATTATCATCATCATTATCGTCATCATCACTA
ATGCTTTTCAACCATACACAAATGTAGGCAGAAGAGTGTAAATCCACTTAGTGAGGCAATCTT
GGAGAGGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAGGACAGTGAGAGGGAA
ATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCATAGTCCACTACATTACTTTG
TATTCTAATATTAAGAAAAATAATAAACCATTCTGTGCACTTATCAOCCAGGCTCAACAG
TTATCTTTGGCCACAGATCCTGTCTCACTGCACTGTCACCTGTCACCTGAGTCCACTTAGCGTTCTG
AATCCAATCCAGGGCATGATGCTTACTCCTACACAGAACTAAAGATTAAAGAGAGTTTAA
AAGTAACCATGACATCTCTCTGTTCTTTAGCGATAAGTTCTTAATATTTATGGCTGCTTGT
GTATGTTCTAATTTCTCTAATATTGTACATTTAGTTGGCAACTACTTTGTTTGAATTGAGT
TGGAGTTAAGGTCCCATAGGATTAATCTCAACATATTTCTATATTATAAACTTTTCTCTCT

Exon 4 ? (Nuc 11286) ↓

TTGTGAAAGTTCTTTGAGAAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATG
AGCTATAACAAAGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGG
TTGCACTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTACGGACTCACTTGC
TCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCATCCAGACACATTGTCTCTTT
GAGCTATCTCCAGATATTTAAAAACCTCCTTCATTAAGNAOCTGTACTTGTATGATAACTTCATTG
ACCTCCCTCCCAAAAGAAATGGTCTCTCTATGCCAAACCTAGAAAAGCCTTTACTTGCATGGAA
ACCCATGGACCTGTGACTGCCATTTAAAGTGGTGTGCGAGTGGATGCAGGGAAACCCAGGTA

Nuc 11680 ↑

ACTATCTGTGTTGTTGTTTCTTTTTTATARKACGTATTTTCTCAATTTTATTAGAAATGA
TATCCCAAAAGTCCCCATAAOCCTCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGG
CCCTGGCATTCCCCTGTACTGGGGCATATAAAGTTTGGGTGTCCAATGGACCTCTCTTCC
AGTGATGGCCAACTAGGCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGG
TACTGGTTAGTTCATAATGTTGTTGCACCTACAGGGTTGAA

Nuc 11967 ↑

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

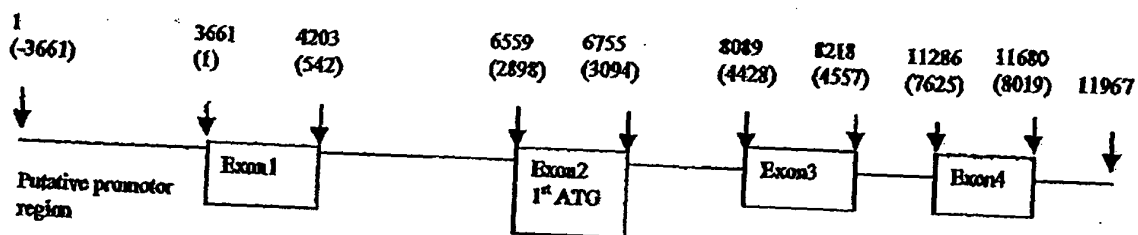


Figure 10

cDNA_rat	1	ogagagacgacagaagggttacggctgcgagaagaogacagaagggtccag
genomic_hu	1	-----
cDNA_rat	51	aaaaaggaaagtgcctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cDNA_rat	101	atgtcacttcagtgaactgagccaggcaaaaogogcggaaggattttgt
genomic_hu	1	-----
cDNA_rat	151	gtagcttgggaocctttcatagacactgatgacacgtttaogaaaaatag
genomic_hu	1	-----
cDNA_rat	201	aaatttgaggagaaaogcctgggecttcggaaaggagtgattgattagta
genomic_hu	1	-----
cDNA_rat	251	cttgcaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu	1	-----
cDNA_rat	301	gggaggaggaagagcattacagagtttccagcagcagcaggaaagctttg
genomic_hu	1	-----
cDNA_rat	351	gttaatttggaaatggatgatagcattaaaataacagaagcgctccagg
genomic_hu	1	-----
cDNA_rat	401	tototgaagottoagtcccccagctgaaagocagaaaagaotaagoccoac
genomic_hu	1	-----
cDNA_rat	451	taagoccttttgatccctttggaagcaagaactttccttccctggggtga
genomic_hu	1	-----
cDNA_rat	501	agactctcctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu	1	-----
cDNA_rat	551	aaaaccaagacagaagagctcaggatgcaggtgagaggcagggaagtcag
genomic_hu	1	-----
cDNA_rat	601	oggettgttgatotoootoaotgotgtotgootgggtgtoacooctggga
genomic_hu	1	-----
cDNA_rat	651	gcagggcotgtocctgcgctgtgocctgatgtgccacagaggtgcac
genomic_hu	1	-----
cDNA_rat	701	tgtacatttcggtacctgacctccatcccagatggcatccoggccaatgt
genomic_hu	1	-----
cDNA_rat	751	ggaacgaataaatttaggatataacagcottaactagattgacagaaaacg
genomic_hu	1	-----
cDNA_rat	801	aotttgatggcctgagcaaaactggagtactcatgotgcacagtaatggc
genomic_hu	1	-----
cDNA_rat	851	attcacagagtcagtgacaagaccttctcgggcttgcagtccttgacaggt
genomic_hu	1	-----
cDNA_rat	901	ottaaaaatgagotataaocaaagtocaaatoattoggaaggataotttot
genomic_hu	1	-----

11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

cDNA_rat	951	acggactcgggagcttggtccggttgcaacctggatcacacaacattgaa
genomic_hu	1	-----
cDNA_rat	1001	ttcatcaaccctgaggccttttatggacttacctogctccgcttggtaca
genomic_hu	1	-----
cDNA_rat	1051	tttagaaggaaaccggtcacaaagctccatccagacacatttgtctcat
genomic_hu	1	-----
cDNA_rat	1101	taagotatctccagatatttaaaacctctttcattaagtacctgttcttg
genomic_hu	1	-----
cDNA_rat	1151	tctgataacttccctgacctccotccocaaaagaatggtotcctacatgcc
genomic_hu	1	-----
cDNA_rat	1201	aaacotagaaagcctgtatttgcattgaaacccatggacctgtgactgcc
genomic_hu	1	-----
cDNA_rat	1251	atttaaagtgggtgtotgagtggtgagcagggaaaocagatataataaaa
genomic_hu	1	-----
cDNA_rat	1301	tgcaagaagaacagaagotcttccagtcctcagcaatgtccctttgcat
genomic_hu	1	-----
cDNA_rat	1351	gaacccacaggatctctaaaggcagaccctttgctatggtacoottgag
genomic_hu	1	-----
cDNA_rat	1401	ctttcctatgtacaaagccaaaccattgatccatcactgaagtoaaagago
genomic_hu	1	-----
cDNA_rat	1451	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt
genomic_hu	1	-----
cDNA_rat	1501	catagaaocotttggtoccttgtctttgaacatgacanannttctggaa
genomic_hu	1	-----
cDNA_rat	1551	ataaggccgacatgggtctgtagtatocaaaagccatcaaggacatcacca
genomic_hu	1	-----
cDNA_rat	1601	actgoattcactgaagaaaatgactacatcatgctaaatgogtoattttc
genomic_hu	1	-----
cDNA_rat	1651	cacaaatottgtgtgcagtgtagattataatcacatccagccagtggtgc
genomic_hu	1	-----
cDNA_rat	1701	aaottctggctttatacagtgactctcctotgatactagaaaggaagccc
genomic_hu	1	-----
cDNA_rat	1751	cagettaccgagactccttcaactgtcttctagatataaacaggtggotot
genomic_hu	1	-----
cDNA_rat	1801	taggocctgaagacatttttaccagcatagaggctgatgtcagagcagacc
genomic_hu	1	-----
cDNA_rat	1851	ctttttggttccaacaagaaaaattgtcttgagotgaacagaaactgcc
genomic_hu	1	-----
cDNA_rat	1901	accacacttagcacattacagatocagttttocactgatgetcaaatcgc
genomic_hu	1	-----
cDNA_rat	1951	tttaccaagggcggagatgagagcggagagactcaaattggaocatgatoc
genomic_hu	1	-----

cDNA_rat	2001	tgatgatgaacaatoccaaaactggaacgcaactgtcctggttggcggcact
genomic_hu	1	-----
cDNA_rat	2051	attgccotgagctgtocaggcaaaaggogaaccttcacctcacttggaaatg
genomic_hu	1	-----
cDNA_rat	2101	gottotagotgatgggagtaaagtgaagagcccccctacgttagogaggatg
genomic_hu	1	-----
cDNA_rat	2151	ggcgaatcctaataagacaaaaatgggaagttggaactgcagatggctgac
genomic_hu	1	-----
cDNA_rat	2201	agctttgatgoaggtotttaccactgcataagcaccaatgatgcagatgc
genomic_hu	1	-----
cDNA_rat	2251	ggatgttotcacatacaggataactgtggtagagccctatggagaaagca
genomic_hu	1	-----
cDNA_rat	2301	cacatgacagtggagtcacgacacagtggttacgggtgagacgctcgao
genomic_hu	1	-----
cDNA_rat	2351	cttccatgcctttccacgggtgttccagatgottotattagotggattct
genomic_hu	1	-----
cDNA_rat	2401	tccagggaaactgtgttctctcagooatcaagagacaggcaaatcttta
genomic_hu	1	-----
cDNA_rat	2451	acaatgggaccttaagaatattacaggttacgcacaaagatoaaggatcat
genomic_hu	1	-----
cDNA_rat	2501	taccaatgtgtggtgccaaacooatoaggggcccacttttccagttttta
genomic_hu	1	-----
cDNA_rat	2551	agtttcagttcaaaagaaaggccaaaggatgggttagcatgcacaggagg
genomic_hu	1	-----
cDNA_rat	2601	caggtggatctggacttggagaacooaactccagtgttcccttaagcag
genomic_hu	1	-----
cDNA_rat	2651	coagcatctttgaaactctotgcacagctttgacagggttoagaggctgg
genomic_hu	1	-----
cDNA_rat	2701	aaaacaagtctccggtgtacataggaagaacaaacatagagacttaatac
genomic_hu	1	-----
cDNA_rat	2751	atcggcggcggtgggattccacgctccggcgattcaggagcataggagg
genomic_hu	1	-----
cDNA_rat	2801	cagctccctctctctgctcggagaattgacccgcaacgctgggcaagcact
genomic_hu	1	-----
cDNA_rat	2851	tctagaaaaagccaaaaagaattctgtgccaaaaagcaagaaaatacca
genomic_hu	1	-----
cDNA_rat	2901	cagtaaagccagtgccactgggtgttccctcgtggaactcactgaogag
genomic_hu	1	-----
cDNA_rat	2951	gaaaaggatgcctctggcatgattcctccagatgaagaattcatgtttct
genomic_hu	1	-----

cDNA_rat	3001	gaaaactaaggcttctggtgtccaggaaggtcaccaactgctgactotg
genomic_hu	1	-----
cDNA_rat	3051	gaccagtaaatcatgggttttatgaogagtatagcttctggcacagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tceactgtgaatocacaaacactaoaatctgagcaccttootgatttcaa
genomic_hu	1	-----
cDNA_rat	3151	attatttagtgtaacaaacgggtacagctgtgacaaagagtatgaacccat
genomic_hu	1	-----
cDNA_rat	3201	coatagcaagcaaaatagaagatacaacaaacccaaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	ttccatcagtagctgaaattcgagattctgtcaggcaggagagagcatc
genomic_hu	1	-----
cDNA_rat	3301	ttccaaagtgcacacootgtaacaggggaaacatggctacctatggco
genomic_hu	1	-----
cDNA_rat	3351	atacacaacatatagtagctttaccagcaagccagtagctcttgag
genomic_hu	1	-----
cDNA_rat	3401	ccaataaatccaaacagaaagttatggacctcagatacctattacaggagt
genomic_hu	1	-----
cDNA_rat	3451	cagcagacotagcagtagtgacatctcttctcacactactgagacccta
genomic_hu	1	-----
cDNA_rat	3501	gcttctcagtcaccccttcaggttcacacaccaactgcotogtotttattt
genomic_hu	1	-----
cDNA_rat	3551	cacattcctagaaaacaacaatacaggttaacttcccttgtccaggcaact
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatttggagcagagggagagttaaaacccacata
genomic_hu	1	-----
cDNA_rat	3651	gaaccccagttctccagggcatagacaaggactgtgaggccagcaatc
genomic_hu	1	-----
cDNA_rat	3701	aagggaacotgctaacaaaatgtgagccaagttccagccaagagtagcca
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttccgagaggggtcaagtggtactg
genomic_hu	1	-----
cDNA_rat	3801	cagcactgtcagttccaagttcatcccaagtgccctcccaaaactaat
genomic_hu	1	-----
cDNA_rat	3851	aatgttgggtcatagcagaagagtctaccactgtggtcaagaaacaaot
genomic_hu	1	-----
cDNA_rat	3901	gttactatttaaggacaaaacaaatgtagatattggataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaaaatattccggaggggaaagtaaccaogtgattoctacggaagca
genomic_hu	1	-----
cDNA_rat	4001	agcatgaottctgtctcaaacatctgtatccotgggaaatctcctgtaga
genomic_hu	1	-----
cDNA_rat	4051	caatagtggtcacctgagcatgcotgggaccatccaaactgggaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtggaaacaacaccacttccagccocctcagacacacctcaataoca
genomic_hu	1	-----

cDNA_rat genomic_hu	4151 1	acaagcacaaaattctcaaagaggaaaactccottgcaacagatctttgt -----
cDNA_rat genomic_hu	4201 1	aaataaccagaagaaggaggggatgttaaagaatccatatcaattoggtt -----
cDNA_rat genomic_hu	4251 1	tacaaaagaaccagcgcaagcttccaaaatagctcctotttttacco -----
cDNA_rat genomic_hu	4301 1	acaggtcagagttccccctcagattotacaactctcttgacaagtcggc -----
cDNA_rat genomic_hu	4351 1	accagctctgtctacaacaatggotgccactcagaacaagggcactgaag -----
cDNA_rat genomic_hu	4401 1	tagtatacaggtgccagaagtctctcagcaggaagaagcagcccttcaco -----
cDNA_rat genomic_hu	4451 1	aactcctctcagtgcttctctagcaccataagcaagagatctaatacatt -----
cDNA_rat genomic_hu	4501 1	aaactttcttgtcaacggaaaacccccacagtgacaagtcctactgtctactg -----
cDNA_rat genomic_hu	4551 1	catctgtcattatgtctgaacccaacgaacaagatccaaagaagcaaaa -----
cDNA_rat genomic_hu	4601 1	gaccaaataaaggggctcgggaagaacagaacaaacgcaaacaccaoooo -----
cDNA_rat genomic_hu	4651 1	caggoaggtttctggctatagtgcatactcagctctaacacacagotgata -----
cDNA_rat genomic_hu	4701 1	cccccttggtctttcagtcattccccacgacaagatgatgggtggaatgta -----
cDNA_rat genomic_hu	4751 1	agtgcagttgcttatcactcaacaacotctcttctggccataactgaact -----
cDNA_rat genomic_hu	4801 1	gtttgagaagtaoaccagactttgggaatacaaacagctttggaacaaa -----
cDNA_rat genomic_hu	4851 1	cgttgttgagcaaatcacaggagagtaaacacagtgaaaagagcctcagac -----
cDNA_rat genomic_hu	4901 1	acacacaccactcctcagcagtgggggcgccccagtgccccactccttc -----
cDNA_rat genomic_hu	4951 1	cccacctccttttactaagggtgtggttacagacagcaagtcacatcag -----
cDNA_rat genomic_hu	5001 1	ctttccagatgaogtcaaatagagtggtcaccatatatgaatcttcaagg -----
cDNA_rat genomic_hu	5051 1	cacaatacagatctgcagcaaccctcagcagagggtagcccaatcctga -----
cDNA_rat genomic_hu	5101 1	gatcataactggaaaccactgactctcctctaatctgtttccatccactt -----
cDNA_rat genomic_hu	5151 1	ctgtgccagcactaagggttagataaaccacagaattctaatggaagccc -----
cDNA_rat genomic_hu	5201 1	tctccctggccagaacacaaatatcagotcaagtcatactccgaaccat -----

cDNA_rat	5251	tgagaagggcaaaaggccagcagtaagcatgtcccccaacctcagccttc
genomic_hu	1	-----
cDNA_rat	5301	cagaggccagcactcatgcctcscactggaatacacagaagcatgcagaa
genomic_hu	1	-----
cDNA_rat	5351	aagagtgttttgataagaaacctgggtcaaaacocaaottccaacatct
genomic_hu	1	-----
cDNA_rat	5401	gccttacgtctctctacctaagactctattgaaaaagccaagaataattg
genomic_hu	1	-----
cDNA_rat	5451	gaggaaagggtgcaagctttacagttccagctaattcagaogtttttctt
genomic_hu	1	-----
cDNA_rat	5501	ccttgtagggotgttgagagccactgcccatcatccactggacccagagt
genomic_hu	1	-----
cDNA_rat	5551	ttcatoaggyanttgaaatatccaaagggacacagaaaagccggttccacg
genomic_hu	1	-----
cDNA_rat	5601	tgottoccaatggcaccttgccatccagagggtcagttatccaggaccgt
genomic_hu	1	-----
cDNA_rat	5651	ggacagtagctgtgtctgcatttaatacactgggcgtagaccattttca
genomic_hu	1	-----
cDNA_rat	5701	tgtototttgtctgtggtttttacccggcaaggattttggacagacatg
genomic_hu	1	-----
cDNA_rat	5751	tcaaggagatccacagttcactttggaagtactgtggaactaaagtgcaga
genomic_hu	1	-----
cDNA_rat	5801	gtggagggtatgccgaggcctacggtttcctggatacttgcaaaccaaac
genomic_hu	1	-----
cDNA_rat	5851	ggtggtctcagaaacggccaagggaagcagaaaggtctggttaacacctg
genomic_hu	1	-----
cDNA_rat	5901	atggascattgatcatotataatctgagtctttatgatcgtggtttttac
genomic_hu	1	-----
cDNA_rat	5951	aagtggtggtggccagcaaccatctggccaggattcactgttggttaagat
genomic_hu	1	-----
cDNA_rat	6001	acaagtcatacagctccctcctgtcattatagagcaaaagaggcaagcca
genomic_hu	1	-----
cDNA_rat	6051	tcgttggggttttaggtggaagtttgaaaotgcctgcactgcaaaaagga
genomic_hu	1	-----
cDNA_rat	6101	actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa
genomic_hu	1	-----
cDNA_rat	6151	accattgcagttgactcattccagattttcttgatccaaatggaactc
genomic_hu	1	-----
cDNA_rat	6201	tgtatataagaagatogotccttcagtgaggggcacttatgagtgcat
genomic_hu	1	-----
cDNA_rat	6251	gccaccagctcctcagggtcagagagaagggtagtgattcttactgtgga
genomic_hu	1	-----

5251
 5301
 5351
 5401
 5451
 5501
 5551
 5601
 5651
 5701
 5751
 5801
 5851
 5901
 5951
 6001
 6051
 6101
 6151
 6201
 6251

rat	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
le_hu	890	CtatGGAAGCAGAATCACAGTCCATAAAATGGAACCTTGGAsattagga
rat	7401	acatccGGCTTTCTGACTCTGCCGACTTCACCTGTGTGGttoggagcgag
le_hu	940	atgtgaGGCTTTCAGATTACAGCCGACTTATCTGTGTGGcccgaaatgaa
rat	7451	ggaGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCCTAGAAATGCTGAG
le_hu	990	ggtGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAG
rat	7501	AAGACCAACATTTCAGAAACCCATTCAACGAAAAagtcatecgcccaagctg
le_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgccagctgg
rat	7551	gcaagcccgTAGCACTGAAGTCTCTGTGGATGGGAACCCCCACCTGAA
le_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTGATGGTAACCCACACCTGAA
rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCAcaaa
le_hu	1140	ataatCTGGATTTTACCAATGGCACAGATTTTCCAATGGACCAcaaaag
rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAgcaa
le_hu	1190	ttatcagTATCTGATAGCAAGCAATGGTTCTTTTATCATTCTAAAcaaa
rat	7701	ctcggaaacagtcAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTGGC
le_hu	1240	ctcgggaggtgCAGGAAAATATCGCTGTGCAGCTAGGAATAAGGTGGC
rat	7751	TAcategagaaactaatctgttagagATTGGGCAGAAGCCAGTCATTCT
le_hu	1290	TAtattgagaaattagtoaatattagaaATTGGCCAGAAGCCAGTTATTCT
rat	7801	GACATACGAACCCAGGGatggtgaagagcgTCAGTGGGGAACCGTTATCAC
le_hu	1340	TACCTATGCACCCAGGGAacagtaaaaggoatCAGTGGAGAATCTCTATCAC
rat	7851	TGCATTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
le_hu	1390	TGCATTGTGTCTGATGGaATCCCTAAGCCAAATATCAAAATGGACTATG
rat	7901	CCGGGTGGccATGTAATCGACAGGCCTCAAGTGGATGGAAAATACATAct
le_hu	1440	CCAAGTGGttATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATatt
rat	7951	GCATGAAAATGGCACGCTGGTCATCAAGCAACAACAGCtcacgacaaAG
le_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCttatgacagAG
rat	8001	GAAATTATATCTGTAGGGCTCAAAACAGTGTGGCCAggcagttattago
le_hu	1540	GAACTATATCTGTAAGGCTCAAAATAGTGTGGTCatacaactgattact
rat	8051	gtgtCAGTGATGGTTGTGGCCTACCCCTCCCCGAATCATAAactacotACC
le_hu	1590	gttcCAGTAATGATTGTAGCCTACCCCTCCCCGAATTACAAAtogtccACC
rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
le_hu	1640	CAGGAGTATTGTCACCAGGACAGGGGCAGCctttCAGCTCCACTGTGTGG
rat	8151	CCTTGGGAATCCCCAAGCCaaAAGTCACCTGGGAGACGCCAAGACACTCC
le_hu	1690	CCTTGGGAGTTCCCCAAGCCagAAATCACATGGGAGATGCCTGACCACTCC
rat	8201	CTGCTCTCAaaagcaacagcaagaaaaccCATAGAAGTGAGATGCTTCA
le_hu	1740	CTTCTCTCAAcggcaagtaagagaggacaCATGGAAGTGAGCAGCTTCA
rat	8251	CCCACAAGGTACGCTgGTCATTGAGAATCTCCAAACCTCGGATTCCGGag
le_hu	1790	CTTACAAGGTACCCTaGTCATTGAGAATCCCCAAACCTCCGATTCTGGga
rat	8301	tcTATAAGTGCAGAGCTCAGAACCTACTTGGgaCTGATTACGCAACAACT
le_hu	1840	taTACAAATGCACAGCAAGAACCACCTTGGtagTGATTATGCAGCAACG
rat	8351	TACATCCAGGTACTCTGACAGGAagggggagactaaaattcaacagaagt
le_hu	1890	TATATTCAAGTAATCTGACATGAaataataaagtcaacaacatctgggca

Figure 12

ATGAAGGTAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTTCAGAAAGATACCTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAATATTGAGTTTATAAACCAGAGGTTTTTATGGGCTCAACTTTCTCCGCTGGTGCCTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCTTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGCCAGATGTAATAAAATGCAAAAAGATAGAAGTCCCTCTAGTGTCTCAGCAGTGT
CCACTTTTGCATGAAACCTTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAG
TGTGCCAAGCCAACCAATTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCT
GCTTTCATCTCTCCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCT
GGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCAATTGCATTCACTGAA
GAAAATGACTACATCGTGCTAAATACCTTCAATTTTCAACATTTTGGTGTGCAACATAGATTACGGTCAC
ATTACAGCCAGTGTGGCAAATTTTGGCTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTG
CTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTACC
AACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAGCTGAAC
AGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGA
GCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACAT
ACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCACCCACACGTGGAT
TGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCCTATGTCAAGTGGATGGACGGATCTTAATAGAC
AAAAGTGGAAAATTTGAACTCCAGATGGCTGATAGTTTGGACACAGGCGTATATCACTGTATAAGCAGC
AATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTGAAGCCTATCAG
GAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACTGGTATC
CCAGATGCCCTCTATTAGCTGGGTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAA
GTTCTAAACAATGGCACATTAAGAATATTACAGGTACCCCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAACCCATCAGGGGTTGATTTTTTGAATTTTCCAAGTTTCAAGTCAAGATGAAAGGACAAAGCCC
TTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCATCCTATTGCTCATCTTAAGGAG
CCACCAGGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACA
AGTAAGAGGCACAACATATCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGG
GAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCCTGTTGGAG
AAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCCACCCCAAGTGGTCACC
CAACTCCCAAACATACCTGGTGAAGAAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATTTATG
GTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG
AGTCCTATGACAAACATAAATTTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCT
GAAGAACCCACAGATTTCAAACGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATAAACCCA
ACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCACTCACTGTCTTTCCACTGCTACTTGGGA
GCAACTGAATTTCAAGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCAAAGTAGACCCCAATA
ACAGTAAGGACTATGATCAAAGATGTCAATGTCAAATGCTTAGTAGCACCACCAACAACTATTATTA
GAGTCAGTAAATACCACAAATAGTCAATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCAC
TTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGCTCAT
TCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTTCAGACGCTTTGGGAGGCAGAGG
AAAATTGGCGGAAGGGGGGGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGC
ATTTTCAGGTCAACAACCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAAT
GTGACATGTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCTTAGTC
CAGAATCCACTATTACTACTTGAGAACAAACCCAGTGATAGAGAAAACAACCCACAATAAAATATTTT
AGGACTGAAATTTCCCAAGTGACTCCAACCTGGTGCAGTCATGACATATGCTCCAACATCCATACCCATG
GAAAAAATCACAAGATAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCA
GTGATTACATCGTCACTTTAGGTGCTATACCAAGCCACCAATGACTATTATAGCCATTACAAGGTTT
TCAAGAAGGAAAATTTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGAAT
CAACATAAAGTTAGTTTACAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCCTGCTTTACCACAG
AGACAAAAGTTCCCTTTTCATTTCACCACACTTTCAACAAGTGTGATGCAAAATTCATCTAATACCTTG
ACTACCGCTCACCACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCC
TTCCACCCCTTAACCTATGCTTCCCTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATCATATCA
ACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCACTTATGAAACCCAAACA
GAGAGATCTAGAGCACAACAATAACAAAGAGAACAGGAGCCTCAAAGAAGAACAGGACTGACCCAAAC
ATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACT

CATTCCCCACCAGAAACACAACTCTGGGATTTCGAAGCAACAATCAGTTTTCATTCAAGAACTCTTAATCTG
ACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCTGAAACA
ACTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCATTAGACACTCCCATACCACCA
TTCTTGAGCAGCAGTGTCTACTCTAATGCCAGTTCCTCATCTCCCTCCCTTTACTCAGAGAGCAGTTACT
GACACACGTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGG
CACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCTCCCTGGGCAGAA
TACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGCATGTTG
GCTACTACAGCGCTGTCCGAGGCCACCACTCTTGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAG
AGTGACTTTGTATAAGAAACCACTTCAAGAAGCAACAACCTCCAAACTCCTTCCCTTTGACTCTTTGTCT
AGGTATATATTTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTACTATTCCAGCTAACTCA
GATGCCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCTGCCACCATTTCATTGGACCAGAGTTTCAGGA
CTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCTGTCCATCCAG
AGGGTGGAAATTGAGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTT
CATGTACCTTGTCTGTGGTTTCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTT
CATTCGGGAAGCACTGTGGAACCTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCTGGATT
CTTGCAAAACCAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGA
ACATTGGTCTCTCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAACCCAGGT
GGCCAGGATTCACTGCTGGTTAAAAATAACAAGTCAATTGCGACACCACCTGTTATTTAGAGCAAAAGGAGG
CAAGTCATTGTAGGCATCTGGGTGAAAGTTTAAATCTGCCCTGTACTGTCAAAAGGAACTCCTCAGCCC
AGCGTTTACTGGGTCTCTCTGATGGCACTGAAGTGAACATTACAGTTTACCAATTCCAAGTTGTTT
TTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAATGCATT
GCTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCC
AGGATAGAAGCTGCATCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAACTGCTCA
GCCACTGGGGAGCCCAAACCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTGCACCAGTGGAGC
TGGATCCACGTCTACCCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTCTAC
TTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTGAAACCT
GCCAAAATTGACCACAAGCAGTATTTTAGAAAAGCAAGTGCTCCATGGGAAAGATTCCAAGTAGATTGC
AAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTGGAGTTTGCCCTGATGGAACCATGATCAACAATGCA
ATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCCTTTTCAACATAGGAATCTTATCTTC
ACAAAGTTTGGGGTACGGGAGGAAGGAGATTATCTGTTATGCCCCAGAACACCTTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAGTAACAAAACCAACAAGAGA
ATCAAAGCTGGAGACACAGCTGTCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGG
TTGCTGCCTTCCAATGACATGATTTCTTTCTCCATTGATAGGTACACATTTTCATGCCAATGGGTCTTTG
ACCATCAACAAAGTGAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCAGAAATCCAGTGGGGAT
GACACCAAAATGTACAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACAAACAGA
ACTGTTATTAAAGCCACAGCTGTGAGACATTTCAAAAAACACTTTGACTGCAGAGCTGAAGGGACACCA
TCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTTCTCTACAGCCCCATACTATGGAAGCAGAATC
ACAGTCCATAAAAATGGAACCTTGGAATTAGGAATGTGAGGCTTTTCAGATTTCAGCCGACTTTATCTGT
GTGGCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGA
CCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCACGCTGGGAAAGTCCACAGCATTTGAATTGC
TCTGTTGATGGTAACCCACCACTGAAATAATCTGGATTTTACCAATGGCACACGATTTTCCAATTGGA
CCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCATAAAACAACTCGGGAGGAT
GCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCA
CTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTATGTAGTA
GACAGGCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAAGCAACA
GCTTATGACAGAGGAACTATATCTGTAAGGCTCAAAATAGTGTTGGTTCATACACTGATTACTGTTCCA
GTAATGATTGTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTACACCAGGACAGGG
GCAGCCTTTCAGCTCCACTGTGTGGCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATGCCTGAC
CACTCCCTTCTCTCAACGGCAAGTAAAGAGGAGACATGGAAGTGAGCAGCTTCACTTACAAGGTACC
CTAGTCATTAGAATCCCAACCTCCGATTCTGGGATATACAAATGCACAGCAAGAACCCTTGGT
AGTGATTATGCAGCAACGTAATTCAAGTAATCTGA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

rat_cdna
human_5+3_corrected
mus_cdna_5

CGAGGACGACAGAGGTTACGGCTGCGAGAAGACGACAGAGGGTCCAGAAAAGGAAA

-----CAGAGGGTCCAGGAAA-GGAAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GTGCTGGAGGGGAGTGGGGACAAAGCAGCGACCAAGTGAATGTCACTTCAGTGAAGTGAAG
GTACTGGAGGGGAGTGGGGACAAAGCAGCGACCAAGGGAACTCGCTTCAGTGAAGTGAAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCCAGGCCAAAAGCGCGGGGAGGATTTTGTGTAGCTTGGGACCTTTTCATAGCACTGAT

 GCCAGGCCAAAAGGAGCGGGGAGGATTTATATGTAGCTGGGAGCTTTTCATAAACACTGAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACACGTTTACGCAAAATA-GAAATTGAGGAGAAACGCTGGGCCTTCGGAAAG---GA
GACGTGTTTGTGCAAAACAGCAATTGAGGAGAAACGCTGGGACGTGGAAAGAAGGA

rat_cDNA
human 5+3 corrected
mus_cDNA 5

GTGATTGATTAGTACTTGCAAGTTTAGGTGACITTAAGGAGAGCTAACTAAATGTATACTA
GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTT--GAGAACTAACTAACTATACTA

rat_cDNA
human 5+3 corrected
mus cDNA 5

TTGAGGGGAGGAGGAAGAGCAATTACAGAGTTTCCAGCAGCAGCAGGAAGCTTTGGTTAAT
TTGAGGGGAGGAAGGARGAGCAITT-----CCAGCAGCAGCAGGAAGCTTTGGTTAGT

rat_cdna
human_5+3_corrected
mus_cdna_5

TTGGAATGGAATGATAGCATTAATAATACAGAAGCGGCTCCAGGTCTCTGAAGCTTCAGT
TTGGAATGTATGATACCATTAATAATACAGAAGCGGCTCCAGTCTCTGAAGAGTTCAGT

rat_cDNA
human 5+3 corrected
mus cDNA 5

CCCCAGCTGAAGCCAGAAAGACTAAGCCCACTAAGCCCTTTGATCCCTTTGGAAGCA
CCCCAGCT-----A-GTGAAGCCTACTAAGCCCTTTGCTCCCGTTGGAAGCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGAACTTTCCTTCCCCTGGGGTGAAGAGCTCTCTCAGAAAGATTTCCTGTCTCTGCCTATG
 AAGAACGTTCTTCAATCAGGTGAAGGCTCTCTCAGAAAGATTTCCTGTCTCTGCTTATG

rat_cDNA
human_5+3_corrected
mus_oDNA_5

TTACAGAGGAATCAAHCCAAGACAGAGAGCTCAGGATCGCGGTGAGAGGCAGGGAAG
 -----ATGAAGGTAAAGGCAGAGCAA
 TTACAGAGGATTCAAAGCAAGACAGAGAGAGCTCAGGATCGCAGAGAGGCAGGGAAG
 *** * * * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

[illegible]

rat_cDNA
human 5+3 corrected
mus_cDNA 5

CCTGTCTCTCGCGCGCTGTGCCTGCTATGTGCCACAGAGGTGCACGTGATACATTTGGGTACC
 CCTGTCTCTCGCGCGCTGTGCCTGTATATGCTAAGGAGGTACACTGCACATTTGGGTACC
 TCTGTCTCTCGCGGATGTGCCTGCTATGTGCCACAGAGGTGCACGTGATACATTTGGGACC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TGACTTCGATCCAGATTGG-CATCCCGGCCAATGTGGAACGATATAATTAGGATATAAC
TGACTTCGATCCAGAGCAG-CATCCCGCCCAATGTGGAACGCATCAATTAGGATACAAC
TGACTTCGATCCAGAGCGGGCATCCAGCGCAATGTGGAACGAGTCAATTAGGGTATAAC
**** * **** * **** * **** * **** * **** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGCCTTACTAGATTGACAGAAAACGACTTTGATGGCTGAGCAAACTGGAGTTACTCATG
AGCTTGCTTAGATTGATGGAAACAGATTTTCTGGCTGACCAAACTGGAGTTACTCATG
AGCCTCACTAGATTGACAGAAAATGACTTTCTGGCTGAGCAGACTGGAGTTACTCATG
*** * ***** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTGCACAGTAATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTCAGTCCCTTG
CTTCACAGCAATGGCATTACAGCAATCCCTGACAGACCTTCTCAGATTTCAGGCTTG
CTGCACAGCAATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTCAGTCCCTTG
** ***** ** ***** * *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCAITCGGAAGGATACTTTCTACGGA
CAGGTCTTAAAAATGAGCTATAATAAAGTCCGAAACCTCAGAAAGATACTTTTATGGC
CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCAITGGAAGGATACTTTGTATGGA
***** ** * ***** ** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCGGGAGCTTGCTCGGCTTGCACCTGGATCACAACACATTGATTCATCAACCTGAG
CTCAGGAGCTTGACAGGATTGACATGGACACAACATATTGAGTTTATTAACCCAGAG
CTCAGGAGCTTGACCGGCTTGACCTGGATCACAACACATTGAGTTTATCAACCCAGAG
*** ***** ** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCTTTTATGGACTTACCTCGCTCGGCTTGGTACATTTAGAAGGAACCGGCTCACAAAG
GTTTTTATGGGCTCAACTTCTCGGCTGGTGCATTTGGAAGGAATCAGCTCACTAAG
GCTTTTACGGACTCACTTGTCTCGGCTGGTACATTTAGAAGGAACCGGCTCACAAAG
* ***** ** * ***** ** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCCATCCAGACACATTGTCTCTTAAGCTATCTCCAGATATTTAAAACTCTTTTCAAT
CTCCACCCAGATACATTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCAAT
CTCCATCCAGACACATTGTCTCTTTGAGCTATCTCCAGATATTTAAAACTCTTTTCAAT
***** ***** ** ***** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGTACCTGTTCTGTCTGATACTTCT-GACCTCCCTCCCAAGAAATGGTCTCCTA
AAGTTCCTATACCTGTCTGATACTTCT-GACCTCCCTCCCTCAAGAGATGGTCTCCTA
AAGNACCTGTACTGTATGATACTTCTATGACCTCCCTCCCAAGAAATGGTCTCCTC
*** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CATGCAAACTAGAAAGCCTGTATTGCAATGGAACCATGGACCTGTGACTGCCATTT
TATGCTGACCTAGACAGCCTTTAAGTGCATGGAACCATGGACCTGTGACTGCCATTT
TATGCAAACTAGAAAGCCTTACTTGCATGGAACCATGGACCTGTGACTGCCATTT
***** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAAGTGGTTGTCTGAGTGGATGCGGGGAAACCATGATATAATAAATGCAAGAAAGACAG
AAAGTGGTTGTCTGAGTGGATGCGGGGAAACCATGATATAATAAATGCAAGAAAGATAG
AAAGTGGTTGTCTGAGTGGATGCGGGGAAACCATGATATAATAAATGCAAGAAAGATAG
***** ** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGCTCTTCCAGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTTAAGGCA
AAGTCCCTGATGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTTAAGGCA
TTCTTTTTTTATARKAGTATTTCTCAATTTCAATTTAGATGATATCCCAAGATC-
* * * * * * * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACCTTTGCTATGTTACCATCTGGAGCTTCTATGTACAAAGCCACCATTTGATCCAT
AGCCGTTAGCTATGGTCTCAGCTGCACTTTCAGTGTGCCAAGCCACCATTTGATCCAT
-CCCATAACTCCCTCCCTCC-CTTCCCTACCTACCATTC-CCATTTTTGGC
** * * * * * * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CACTGAAGTCAAGAGCCTGTTTACTCAGGAGGACAAATGGATGTGCTCCACCTCACTC
CCTGAAATCAAGAGCCTGACTATTCTGGAAGACAGTAGTCTGCTTTCTATCTCTCCCT
CCTGGCAATCCCT-
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGATTTATAGAAACCTTTGGCTCTTGTCTTTGAACATGACANANTNTCTGGAATA
AAGTTTCAATGGCAACCTTTGGCTCTTGTCTTTGAATATGACAGATCATCTGGAATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGGCCGACATGGTCTGTAGTATCCAAAGCCATCAAGGACATCAACCACTGCATTCACTG
AAGCTAACATGGTCTGCACTATTCAAAAGCCCTCAAGGACATCAACCACTGCATTCACTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAGAAATGACTACATCATGCTAAATGCGTCATTTCCACAAATCTTGTGTGCAGTGTAG
AAGAAATGACTACATCGTCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATTATAATCACATCCAGCCAGTGTGGCACTTCTGGCTTTATACAGTGACTCTCTCTGA
ATTACGGTCACATTCAGCCAGTGTGGCAATTTTGGCTTTGTACAGTGATTTCTCTCTGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TACTAGAAAGGAAGCCCCAGCTTACCGAGACTCCTTCAGTGTCTTCTAGATATAACAGG
TACTAGAAAGGAGCCACTTGTCTAGTGAACACCCGAGCTCTATTACAATATAAACAGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGCTCTTAGGCGTGAAGACATTTTACCAGCATAGAGGCTGATGTGAGAGCAGACCOCTT
TGGCTCTTAGGCGTGAAGACATTTTACCAGCATAGAGGCGATCTCAGAGCAGATCOCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTGGTTCCACAAAGAAAAATTTGTCTTGAGCTGAACAGAACTGCCACCACTTAGCA
CTTGGTTAATGCAAGACCAATTTCTTGAGCTGAACAGAACTGCCACCACTTAGCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CATTACAGATCCAGTTTTCAGTGTGCTCAAAATGCTTTACCAAGGGCGAGATGAGAG
CATTACAGATCCAGTACTCCAGTGTGCTCAAAATCATTTCACAGAGCAGAGATGAGGC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGAGAGAGCTCAAATGGACATGATCTGTGATGAACATCCAACTGGAAAGCAGCTG
CAGTGAACACAAATGGACTATGATTTCAAGGGATAACATCTAGCTGGAACTACTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TCTTGGTTGGCGCCTATTGGCTGAGCTGTCCAGGCAAGGGGACCOCTTCACTCTACT
TCTTGGTTAGGTGAACCGTTGGCTGAGCTGCCAGGCCAAGGAGACCCACCCCTACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTACGTTAGCGAGGATGGC
TGGATTGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTATGTCAGTGAGGATGGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GAATCCTATAGACAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGATGCAG
GGATCCTATAGACAAAATGGGAAGTTGGAAGTGCAGATGGCTGATAGTTTGTACACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GTCTTTACCACTGCATAAGCAACCAATGATGCAGATGGGATGTTCTCACATACAGGATAA
GGGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTGGTAGAGCCCTATGGAGAAAGCACATGACAGTGGAGTCCAGCACAGTGGTTA
CTGTGGTAGAAGCTTTGGTGGAGCCCTATCGGAAATGGGATTCATCACAGTGGTTA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGGGTGAGAGGCTCGAAGCTTCAATGCTTTCCAGGGGTGTTCCAGATGCTTCTATTAGCT
TTGGTGAACACTTGATCTTCCATGCAATCTACTGGTATCCAGATGCTCTATTAGCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAGACAGGCAATTTCTAACA
GGGTTATTCCAGGAACCAATGTGCTCTATCAGTCATCAGAGACAGAAAGTTCTAACA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATGGGACCTTAAGAAATATTACAGGTACGCCAAAGATCAAGGTCATTACCAATGTGTGG
ATGGCACATTAAAGAAATATTACAGGTACCCCCGAAGACCAAGGTTATTATOGCTGTGTGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGCCAACCCATCAGGGGGCGACTTTTCCAGTTTAAAGTTTCAGTTCAAAGAAAGGC
CAGCCAACCCATCAGGGGTTGATTTTGTATTTCCAGTTTCAGTCAAGATGAAGGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAAGGATGGTTGAGCATGACAGGGAGGCGAGTGGATCTGGACTTGAGAACCCAACTCCA
AAAGGCCCTTGGAGCATGATGGAGAACAGAGGGATCTGGACTTGATGATCCAATCCTA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGTGTTCCCTTAAGCAGCCAGCATCTTTGAACTCTCTGCATCAGCTTTGACAGGGTCAG
TTGCTCATCTTAAGGAGCCACAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGGCTGGAAAACAAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC
AGGTTGGAAAACACACCTCAAGCACAGTAAGAGGCACAACTATCGGGATTAACACTCC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGCGGCGTGGGGATTCCACGCTCGGGCGATTCAAGGAGCATAGGAGGCAGCTCCCTCTCT
AGCGAOGTGGAGATTCAACACATCGACGTTTAGGGGGAATAGGAGGCATTTCCTCCCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGCTCGGAGAAATTGACCCGCAACGCTGGGCGCACTTCTAGAAAAGCCAAAAGAAATT
CTGCTAGGAGAAATTGACCCACAACTTGGGCGCACTGTTGGAGAAAGCTAAAAGAAATG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTGCCAAAAGCAAGAAATATCCACAGTAAGCCAGTGCCACTGGCTGTTCCCTCG
CTATGCCAGACAAGCGAGAAATATCCACAGTGAGCCACCCTCAGTGGTCAOCCAACTCC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGAACCTCACTGACGAGGAAAGGATGCTCTGGCATGATTCCTCCAGATGAAGAATTCA
CAAACATACTGCTGGAAGAGACGATTCCTCAGGCATGCTGCTCTACATGAGGAATTIA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGTTCTGAAACTAAGGCTTCTGCTGTCCAGGAGGTCAACCACTGCTGACTCTGGAC
TGGTCCCGGCCACTAAAGCTTTGAACCTTCAGCAAGGACAGTGACTGCTGACTCCAGAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CAGTAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAGTCTCAACTGTGAATC
CAATATCTGATAGTCTATGACAAACATAAATTATGGCACAGAACTCTCCGTTGTGAATT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CACAAACACTACATCTGAGCACTTCTGATTTCAAAATATTTAGTGAACAAAGGTA
CACAAATFACTACCACTGAAGAACCAAGATTTCAAACTGTCTACTGCTATTAACATA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CAGCTGTGACAAAGAGTATGAACCATCCATAGCAGCAAAATAGAGATACACCAACC
CAGCCATGTCAAAGATATTAACCCAAOCTGTCAAGCCAAATACAAGGCACAAOCTATC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAAACCCATCATTAATCTTTCCATC-----AGTAGCTGAATTCGAGATTCTG-CT
AACATTCTATCATCTGTCTTTCCACTGCTACTTGGAGCACTGAATTTAGGACTCTGACA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGCA---GGAAGAGCATCTTCCAAAGTGCAACCCCTGTAAACAGGGGAAACATGGCT
GAGGGAAGAGGAAGAGGACATTTCC--AGTA-ACCCCAATAACAGTAAGGACTATGATC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACCTATGGCCTAACACACATA--TAGTAGCTTTACCAGCAAAGCCAGTACAGTCTTGC
AAGATGNTCAATGTCAAANATGCTTAGTAGCAACCAACAAA-CTATTA-----TTAG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGCCATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCAGCAGAC
AGTCAGTAAATACCAAAATAGTCAT-----CAGACATCTGTAAAGAGTAAGTGAAC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCTAGCTTCTCCAGTCACCCCTT
CCAGGCACAATCACTTCTATTCTCACACTACTCAAATACTTAGCACCTCCAGCTTCCCTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGTTCCACACACCCTGCTCTTTATTTCACATTCTAGAAACAACATACAGGTA
CAGATCCACACACAGCTGCTCATTTCTCAGTTTCCGATCCCTAGANNNAATAGTACAGTTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACTTCCCTTGTCTCAGGCACTTGGGAAGAGAGAGGACAAITTTGGAGCAGAGGGAGAGTTA
ACATCCCGCTGTTCAGACGCTTGGGAGGCAGAGGAAATTTGGCGGAGGGGGGAGTTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAAACCCACATAGAAACCCAGTTCTCCGACGGCATAGACACAGGACTGTGAGGCCAGCAA
TCAGCCCATATAGAACTCCAGTTCTGCGAAGGCATAGATACAGCATTTTCAGGTCACAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TTAAGGGACCTGCTAACAAAATGTGAGCCAGTTCCAGCCACAGAGTACCTGGGATGT
CCAGAGGTTCTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCCACACATGTCTTCCGAGAGGGGCTCAGTGGCTACTGCAGCACTGTCACTTCCAA
GTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTTGTCTTTCCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTTCATCCACAGTGGCTCCCAAACTAATAATGTTGGGGTCATAGCAGAAGAGTCTA
GTGCTGTCTCCATCACCTTCCCAAGGCTGACATTGCTAGAGTCCCATCAGAGAGTCTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCACTGTGGTCAAGAAACACTGTTACTATTAAAGGACAAACAAATGTAGATATTGAGA
CAACTCTAGTCCAGATCCACTATTACTTCTGAGAACAAACCCAGTGTAGANNNGAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TAATAACAACCACTACAAAATATTCCGGAGGGGAAGTAACCAAGTGATTCCTACGGAG
NNACAACACCCACAATAAATAATTCAGGACTNGAAATTTCCCAAGTGACTCCAACCTGGTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAAGCATGACTTCTGCTCCACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTG
CAGTCATGACATATGCTCCACATCCATACCCATGGGAAAACTCACAAGTAACGCCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTCACCTGAGCATGCTGGGACCATCCAACTGGGAAAGATTCAAGTGGAAACACACCAC
GTTACCCACGTGTGTCTAGCAACCAATGAAGCTAAAAGAGATTCAAGTGATTACATGTCAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTCCAGGCCCCCTCAGCACACCTCAATACCA-----CAAGCACAAATTCCTAAAGA
TTTCAGGTGCTATCAOCCAAGCCCAATGCTATTATAGCCATTACAGGTTTTCAGAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGAAACTCCCTTGACACAGATCTTTGTAATAACAGAGAGAGGGGATGTTAAGA
GGAAATTCCTGGCAACAGACTTTGTAATAACCAACCAAAAGGCAGATTAAAGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATCCATATCAATTGGTTTACAAAAGAACCCAGCCGCAAGGTTCCCAAAATAGCTCTC
ATCAACATAAGTTAGTTTACAAAAGCCACAGCTGTGATGCTTCTCTAAACATCTCTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTTACCCACAG-CTCAGAGTTCCOCTCAGATTCTACAACTCTCTTGACAAGTCCSOCA
CTTTACC-ACAGAGACAAAGTTCCOCTTTCCATTTCACCACTTTCAACAGTGTGATG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCAGCTCTGTCTACAACATGGCTGCCACTCAGAACAGGGCACTGAAGTAGTATCAGGT
CAAATCCATCTAATACTGACTACCGCTCACCACACTACGACAAA--ACACACAA-T

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GCCAGAGTCTCTCAGCAGGGAAGAACGCCCTTCAOCCAACTCTCTCCAG-TGCTTCC
CCTGGAAAGTCTTCCACAAAGAGGAGCTTCCCTTC-CCACCCCTTACCCCTATGCTTCC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TAGCOCATAAGCAAGAGATCTAATACATTAACTTCTTGTCAACGGAAACCCACAGT
TAGTATTATAAGCAAGACTCAAGTACAAAAGCATCATATCAACGCCAAACAGCAACGC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GACAAGTCTACTGCTACTGCTATGTCTGTAACCCACGAACAGATCCRA
AACAACTCTACTCTTCCCTGCTATGTCTACTTATGAACCCAAACAGAGAGATCTAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGAGGCAAAAGACCAATAAAGGGGCTCG--GAAGAACAGAAACAGCAACACAC
AGCACAAACAATAAAGAGAGGAOCTCAAAAGAGAACAGGACTGACCCAAACATCTC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCCCAGGCAGGTTTCTGGCTATAGTGCATCTCAGCTCTAACACAGCTGATAOCCCTT
TCCAGACCAAGATTCTGGCTTCACTACCCACTGCTATGACNAGCTCTGCTCTNN

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGCTTTCAGTCAATCCOCCAGACAGATGATGGTGGAAATGTAAGTGCAGTTGCTTATCA
NGCATTCACCTCAATCCOCCAGAAACACACACTGGGATTTCAAGCACATCAGTTTCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTCAACAACCTCTCTCTGGCCA---TACTGAACTGTTTGG-AGGTAC--AOCAGAC
TTCAAGACTCTTAATCTGACAGATGTGATTGAGAGACTAGCCACAGTACTCTAGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTGGGAATACACAGCTTTGGAACAACTTGTGAGCAATCACAGGAGAGTACAC
TTTGAAGAGCACAAATGCTTCTGAAACAACTTGTCCAGCAATCACACAGAGTACAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGTGAAGAGGCTCA--GACAC--ACCAACCACTCTCAGCAGTGGGGGCCCCC
AACTAGGAAGCATATTAGACACTCAACCAACCACTTCTTGAAGCAGCTGCTACTCT

rat_cdna
human_5+3_corrected
mus_cdna_5

AGTGCOCACCTCCTTCCACCTCCTTTTACTAAGGGTGTGGTTACAGACAGCAAAGTCAC
AATGCCAGTTCCCATCTCCCTCCCTTTACTCAGAGAGCAGTTACTGACAAAGTGGGAC

rat_cdna
human_5+3_corrected
mus_cdna_5

ATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA
TOCCATTTCOGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCYCAAGGCACAA

rat_cdna
human_5+3_corrected
mus_cdna_5

TACAGATCTGCAGCAACCTCAGCAGAGGCTAGCCCAATCCTGAGATCATAACTGGAAC
TOCMNNNNNNCA--AATGCCAAGTTCA-----CNAATTTGNGAACNNNNNACTCRRNN

rat_cdna
human_5+3_corrected
mus_cdna_5

CACAGACTCTCCTCTAACTCTGTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATAA
NACTTCATCTACNTCTAACTCTGTACATTCTACTCCCATGCCAGCACTAACACAGTTAA

rat_cdna
human_5+3_corrected
mus_cdna_5

ACCACAGAATTCTAAATGGAAGCCCTCTCCTCGGCCAGACACAAATATCAGCTCAAGTC
ATCAGAGAATTCCAAATTAAGTCCATCTCCTCGGGCAGAAATACCAATTTGGCAGAAACC

rat_cdna
human_5+3_corrected
mus_cdna_5

ATACTCGGAAACCATTTGAGAGGGGCAAAAGGCCAGCAGTAAGCATGTCCCCCACTCAG
ATACTCAGACATTGCTGAAAAGGCCAAAAGGCCAGAGTAAGCATGTTGGCT-ACTACAG

rat_cdna
human_5+3_corrected
mus_cdna_5

-CCTTCAGAGGCCAGCACTCTATGCTCAGACTGGAATACACAGAAGCATGCAGAAAGA
GCTGTCCGAGGCCCACTCTTGTTCAGATTGGGATGGACAGAGACACAAAGAGAGA

rat_cdna
human_5+3_corrected
mus_cdna_5

GTGTTTTTGATAAGAAACCTGGTCA--AACC--CAACTCCAAACATCTGCTTACGTCT
GTGACTTTGATAAGAAACAGTTCCAGAGGACAGACTTCCAAACTCCTTCCCTTTGACT

rat_cdna
human_5+3_corrected
mus_cdna_5

CTCTACCTAAGACTCTATTGAAAAAGCCAGATAATTGGAGGAAGGCTCCAGCTTTA
CTTTGTCTAGGTATATATTGAAAAGGCCAGGATAGTTGGAGGAAGAGCTCCAGTTTAA

rat_cdna
human_5+3_corrected
mus_cdna_5

CAGTTCAGCTAATTCCAGACGTTTTCTCTCTGTGAGGCTGTGGAGAACCACTGCCCA
CTATTCCAGCTAATCTCAGATGCTTTCTTCCCTGGAAGCTGTGGAAATCCCTGCCCA

rat_cdna
human_5+3_corrected
mus_cdna_5

TCATCCTACTGGACAGAGTTTCATCAGGANTTGAATATCCAGAGGACACAGAAAAGCC
CCATTCAATTGGACAGAGTNNNNTCAGGACTTGATTATCTAAGAGGAACAGAAATAGCA

rat_cdna
human_5+3_corrected
mus_cdna_5

GGTTCACGTGCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTG
GGGTCCAGGTTCTCCCAATGGTACCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCG

rat_cdna
human_5+3_corrected
mus_cdna_5

GACAGTACCTGTGCTCTGCATTAAATCAGTGGGCTAGAACATTTTCAGTCTCTTTGT
GACAGTACTTGTGTTCCGCATCCAACTCTTTTGGCACAGACCACTTCATGTACCTTGT

rat_cdna
human_5+3_corrected
mus_cdna_5

CTGTGGTTTTTTTACCCGCAAGGATTTTGGACAGACATGTCAAGGAGATCAGTTTCACT
CTGTGGTTTCTATCTCTCCAGGATCCTGGAGAGACGTACCAAGAGATCACAGTTCAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TTGGAGTACTGTGGAACTAAAGTGCAGAGTGGAGGSTATGCGAGGCTACGGTTTCCT
CCGGAGCACTGTGGAACTGAGTGCAGAGCAGAGGTAGGCCAAGCCCTACAGTTAOCCT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGTACTTGCAAACCAACGGTGGTCTCAGAACGGCCAGGGGAAGCAGAAAGGTCTGGG
GGATTCTTGCAAACCAACAGTTGTCTCAGATCATCCAGGGAGTAGGCAGGCTGTGG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTACA
TGACGGTTGACGGAACATTGGTCTCCACAACTCTCAGTATTATGACCGTGGCTTTACA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTGGTTAAGATACAGTCATCA
AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCAGTGTGGTTAAGATACAGTCATTC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGCTCCOCCTGTCTATTATAGAGCAAAAGAGGCAAGCCATCGTTGGGGTTTTAGGTGGAA
CAGCACACCTGTTATTCTAGAGCAAAAGAGGCCAGTCATTGTAGGCACCTTGGGGTGAAG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTTTGAAACTGCCCTGCACTGCAAAAGGAAGTCCCGAGCCTAGTGTTCAGTGGGTCTTT
GTTTAAAACTGCCCTGTACTGCAAAAGGAAGTCCCGAGCCTAGTGTTCAGTGGGTCTTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATGATGGGACTGAACATAAAACCAATTGCAGTTGACTCAITCCAGATTTTCTGTATCCAA
CTGATGGCACTGAAGTGAAACCAATTACAGTTTACCAATTCAGATTTGTTCTTATTTTCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATGGGACTCTGTATATAAGAGCATCGCTCCTTCAGTGGGGGCATATAGAGTGCAITG
ATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCATATAGAGTGCAITG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCACAGCTCCTCAGGCTCAGAGAGAGGGTAGTGATTCTTACTGTGGAGAGGGAGAGA
CTACAGTTCACATGGTTTCGGAGGAGAGTAGTATGCTTACAAATGGAAGAGGAGTAGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAATCCCGAGTAGAAGCTGCTCTCAGAAATGGAGTGAATTTGGGTGAGAAAT
CCAGCCCGAGTAGAAGCTGCTCTCAGAAATGGAGTGAATTTGGGTGAGAAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TACTACTGAAGTGTCTCAGTACTGGGGATCCAAAGCTAGATAATCTGGAGGCTGCCAT
TACTACTGAAGTGTCTCAGTACTGGGGATCCAAAGCTAGATAATCTGGAGGCTGCCAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCAGGCTGTCTGACACAGTGGCCAGAAATGGGAGCCGAATCCAGCTCTACCCAAATG
CCAGGCTGTGTGACACAGTGGCCAGAAATGGGAGCCGAATCCAGCTCTACCCAAATG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GATCCTTGGTGGTTGGGTGAGTACGGGAAAAAGACGCTGGTGAATCTATGTGTGGCAA
GATCCTTGGTGGTTGGGTGAGTACGGGAAAAAGACGCTGGTGAATCTATGTGTGGCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GAACAAAAATGGGAGATGACCTAGTCTGATGATGTCCGCTGAGATTGACACCTGCCA
GAACAAAAATGGGAGATGATCTGATGATGTCCGCTGAGATTGACACCTGCCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAATTGAACAGAAGCAGTATTTTAAAGCAAGTGTCCATGGGAAAGATTTCAGTGTG
AAATTGACCAACAGCAGTATTTTAGAAGCAGTGTCCATGGGAAAGATTTCAGTAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCTGATGGGACAG
ATTGCAAGCTTCCGGCTCCCACTGCCAGAGATATCTTGGAGTTTGCTGATGGAAOCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT
TGATCAACAATGCAATGCAAGCGATGACAGTGGCCAGGACTAGGAGATATACCTTT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TCACAAATGGAACTTGTATTTCACCAAGCTTGGGATGGCAGAGGAAGGAGATTATCT
TCAACAATGGAACTTTATACCTTCAACAAGTTGGGTAGGGAGGAAGGAGATTATCTT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GCTCTGCCAGAACCTTAGGGAAAGATGAGATGAATGTCACCTAACAGTTCTAACAG
GCTATGCCAGAACCTTAGGGAAAGATGAATGAAGGTCCACTTAACAGTTATAACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCATCCACGGATAAGGCAAGCTACAGACCAACATGAGGCTCAGGGCTGGAGAACAG
CTGCTCCCGGATAAGGCAGAGTAACAAACCAACAGAGAAATCAAAGCTGGAGACACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTATTTGGTTGCTGCCTT
CTGTCTTGACTGTGAGGTCACTGGGGAACCGAAGCCCAATGTATTTGGTTGCTGCCTT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCAACAATGTCAATTCATTCTCAATGACAGGTTCACATTTCAAGCAATAGACTTTGT
CCAATGACATGATTTCTCTTCTCAATGATAGGTACACATTTCAAGCAATGGGTCTTTGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CCATCCATAAAGTGAAACCACTTGACTCTGGGACTATGTGTGCGTAGCTCAGATCCTA
CCATCAACAAGTGAAACTGCTCGATTCTGGAGAGTAAGTATGTGTAGCCGAAATCCCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GTGGGGATGACACTAAGACATACAACTGGACATTTCTCTTAACCTCCATTATCAATG
GTGGGGATGACACCAAAATGTACAACTGGATGTGTCTCTAAGCTCCATTATCAATG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GCCTGTATGCRAACAAGACTGTTATTAAAGCCACAGCCATTCGGCACTCCAAAAATACT
GTCTGTATACRAACAGACTGTTATTAAAGCCACAGCTGTGAGACATTCAAAAAAGACT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTGACTGCAGAGCAGATGGCATCCATCTTCCAGGTCACTGGATTATGCCAGGCATA
TTGACTGCAGAGCTGAAGGGACACCATCTCTGAAGTCATGTGATCATGCCAGCAATA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTTCTCCAGCTCCATACTTTTGGAGCAGAGTCAGGTCATCCAAATGGAACCTTGG
TTTTCTCCAGCTCCATACTATGGAAGCAGAAATCAAGTCCATAAATGGAACCTTGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AGATGAGGAACATCCGCTTTCTGACTCTGCGGACTTCACTGTGTGGTTCGGAGCGAGG
AAATTAGGAATGTGAGGCTTTCAGATTCAAGCGACTTTATCTGTGTGGCCAAATGAGG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GAGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT
GTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCAGAAACCCATTCAACGAAAAAGTCATGGCCAGCTGGCAAGCCCGTAGCACTGAACT
TTAGAAATCCATTTAATGAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCAATTGAATT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCTCTGTGGATGGGAACCCCACTGAAATTAAGTGGATCTTACCTGACGGCACACAGT
GCTCTGTGTATGGTAACCCACACCTGAAATAATCTGGATTTTACCAATGGCACACGAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TTGCTAACAGACCCACAAATCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCTTT
TTCCCAATGGACCAACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTATCATTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACAAAGCAACTGGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
CTAAACAACTCGGGAGGATGCAGGAAATATCGCTGTGCAGCTAGGAATAAGGTTGGCT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACATCGAGAACTCATCTGTAGAGATTGGGCAGAGCCAGTCATTCTGACATACGAAC
ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAGCCAGTTATTCTTACCTATGCAC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGATGGTGAAGAGCGTCAGTGGGGAACGTTATCACTGCATTGTGTCTCTGATGGGA
CAGGACAGTAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTCTCTGATGGGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCCCCAAGCCAAATGTCRAAGTGGACTACACCGGGTGGCCATGTAATGGACAGGCTCAAG
TCCCAAGCCAAATATCAAAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCTCAAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TGGATGGAAATACATACTGCTGAAATGGCACGCTGGTCATCAAGCAACACAGCTC
TTAATGGGAATACATATTGCATGACAAATGGCACTTAGTCATTAAAGAGCAACAGCTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACGACCAAGGAATATATCTGTAGGGCTCAAAACAGTGTGGCCAGGCACTTATAGCG
ATGACAGAGGAACATATCTGTAGGGCTCAAAATAGTGTGGTCATACACTGATTACTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TGTCAGTGTATGGTTGTGGCCACCCCTCCCGAATCATAACTAOCCTACCCAGGACATGC
TTCCAGTAAATGATTGTAGCCTACCCCTCCCGAATTACAAATGTCACCCAGGAGTATTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCCCTGGGAAATCCCAAGCCAA
TCACAGGACAGGGGAGCCCTTTCAGCTCCACTGTGTGGCCCTGGGAAATCCCAAGCCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAGTCACCTGGGAGACGCGACAGCACTCCCTGCTCTCAAAAGCAACAGCAGAAACCCC
AAATCACATGGGAGATGCTGACCACTCCCTTCTCTCAACGGCAAGTAAGAGAGGAGCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATAGAAGTGAGATGCTTACCCACAGGTACGCTGGTCATTCAAGATCTCCAAACCTGG
ATGGAAGTGAGCAGCTTCACTTACAGGTACCCCTAGTCATTCAAGATCCCAACCTGG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATTCCGGAGTCTATAAGTGCAGAGCTCAGAACCTACTTGGGACTGATTACGCAACACTT
ATTCTGGGATATACAAATGCACAGCAAGAAACCCACTTGGTAGTGATTATGCAGCAACGT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACATCCGGTACTCTGACAGGAGGGGGAGACTAAAATTACAGAGTCCACATCCACA
ATATTCAAGTAATCTGACATGAA-----ATAATAAGT-CAACAA-----CATCTGGGCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGGTTTATTTTGGGAAGAAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGA
GAATTTATTTTGGGAAGAAGTTTAAATCAAAGGCAGCCATAGGCATGTAAATGAATTTGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATACATTACAGTATTAAATTTACAATGGACATGCGA--TGA--GACTTGTAATGAAA
ATACATTACAGTATTAAATTTACAATGACATGCAAAATAAAGGACTTGTAAATTAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCATTGTGAACTGAA--ACCGAGTCTCTG--TGGATCTCAAAGCCTTCTTAACTTAA
GCATTATGAACGTATGATCTGATTTATTTAATGGATCTCAAACAAACTTTTAACTTAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGCACTTTGATTTTGGCAACAAATAATAACAAACATTAAGAGAAAAAATGATCCACTAC
GGCACTTTTATTTTGGCAACAAATAATAACAA-----ACATTGAAACGGTTCACTAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GAAATAACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAAGACCTTCTCTCGCTAAC
AAAATAACAAATGGCTAATGTACCTGAATTTTTCAGTAAAAAA--TGAACCTT-CTAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGTTGCCAGCTGCTCTGTCTCTTCTACCAATGTCAAAACATCGCACACAGGGTGA
A-----CCAGTTGCCTAGTGTCCACTCTCTATCAATGTACAAAGCATGGCACTCAG-----A

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATGGAGTCAACGGGAAGATTAAAGTTTGGGTCGTGTAAATCTCAATGTACAAATATTC
ACAGAGACAAATGGAAAAATTAATCTGC-----AATCT--ATGTATAAATATTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TGTCTCTGGTTTATAAACATTTT-GATAAAACCGAAAAAATAAATAAATAAATAAATAA
TGT-----GGTTTATAAATTTTTTGGCTAAACCTACAGAAATAAG-----

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAAA

(rat_cDNA: SEQ ID NO:7)
(human_5+3 corrected: SEQ ID NO:8)
(mus_cDNA_5: SEQ ID NO:9)

Figure 15

rat	MQVREGREVSGLLISLTAVCLVVTGSRACPRRCACVPTVEVHCTFRYLTSIPDGIPANVE
human_5+3_corrected	MKVGRGITCLLVSEFAVICLVATPGGEACPRRCACVPTVEVHCTFRYLTSIPDSIPPVNE
mouse_5_corrected	MOKRGREVSCLLISLTATCLVVTGSRVCPRCACVPTVEVHCTFRDLTSIPD-GPANVE
rat	RINLGYNLSLTRLTENDFGLSKLELIMLHSHNGIHRVSDKTFSGLSQSLQVLKMSYKVKQII
human_5+3_corrected	RINLGYNLSLVRIETDFSGLTLELIMLHSHNGIHTIPDKTFSDLQALQVLKMSYKVRKL
mouse_5_corrected	RVNLGYNLSLTRLTENDFSGLSKLELIMLHSHNGIHRVSDKTFSGLSQSLQVLKMSYKVKQII
rat	RKDTFTGLGSLVRLHLDENNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
human_5+3_corrected	QKDTFYGLRSLTRLHLDENNIEFINPEVYGLNTRLVHLEGNQLTKLHPDTFVSLSYLQ
mouse_5_corrected	EKDTLYGLRSLTRLHLDENNIEFINPEAFYGLTLRLVHLEGNRLTKLHPDTFVSLSYLQ
rat	IFKTSFIKYLELSDNFLTSLPKENVSYMPNLESYLYHGNPWTCDCHLKNLSEWMOGNPDI
human_5+3_corrected	IFKISFIKFLYLSINFLTSLPQENVSYMPDLDSLYLEGNPWTCDCHLKNLSDWIQ--PDV
mouse_5_corrected	IFKTSFIKLYLYDNF-TSLPKENVSSMPNLESYLYHGNPWTCDCHLKNLSEWMOGNP--
rat	IKCKKDRSSSSPQQCPLOMNPRI SKGRPFAMVPSGAFCTKPTIDPSLKSKSLVTOEDNG
human_5+3_corrected	IKCKKDRSPSSAQCCPLCHNPRTSKGRPLAMVSAAPQCAKPTIDSLKSKSLTILEDS
mouse_5_corrected	-----
rat	SASTSPQDFIEPFGSLSLNMTXISGNKADMVCSIQKPSRTSPTAFTEENDYIMNASFST
human_5+3_corrected	SAFISPGGFMAPFGSLTIANMTDQSGNARMVCSIQKPSRTSPAFTEENDYIVLNTSFST
mouse_5_corrected	-----
rat	NLVCSVDYNIQPVWQILALYSDSPLILERSKQLTETPSLSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIUYGHIQPVWQILALYSDSPLILERSHLLSETPOLYYKYKQVAPKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADFFWFQCKIVLQNRATATLSTLQIQFSTDAQIALPRAEMRAERLKWTHIMLNN
human_5+3_corrected	DLRADPSWIMQDQISLQNRATATFSTLQIQYSSDAQITLPRAMRPVKHKNTHISDNN
mouse_5_corrected	-----
rat	PKLERTVLVGGTIALSCPGNGDPSPHLENLLADGSKVRAPYVSEDGRILIDKNGKLELQ
human_5+3_corrected	TKLEHTVLVGGTVGLNCPGQDPTPHVDWLLADGSKVRAPYVSEDGRILIDKNGKLELQ
mouse_5_corrected	-----
rat	ADSFDAGLYHCISTNDADADVLTIRITVVEFYGESTHDSGVQHTVVIGETLULPCLSTGV
human_5+3_corrected	ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIIHHTVFIGETLULPCHSTGI
mouse_5_corrected	-----
rat	PDASISWILEPNTVFPSPSRDRQILNNGTLRILOVTPKQGHYQCVAAAPSGADFSFKV
human_5+3_corrected	PDASISWVTPGNVLYQSRDKVILNNGTLRILOVTPKQGGYTRCVAANPSGVDFLIPQV
mouse_5_corrected	-----
rat	SVQKKGQRMVEHREAGGSGLGEPSN8SVSLKQPAKLSASALTGSEAGQVSGVHRKNK
human_5+3_corrected	SVKMGQRFLEHDGETEGSGLDESNIPIAHKEPPGAQLRTSALMEAEVGGKHTGSTKRRN
mouse_5_corrected	-----
rat	HRDLIHRRGDSTLRRFREHRRQLPLSARRIDPQWAALEKAKKNSVPPKQENTTVKPV
human_5+3_corrected	YRELTLORRGUSTHRRFRENRHFPFSARRIDPQWAALEKAKKQAMPKRENTTVSPF
mouse_5_corrected	-----

rat PLAVPLVELTDEEKDASGMIPDEEFMVLTKASGVPRSPADSGPVNHGEMTSIASGT
human_5+3_corrected PVVTQLENI PGEE DOSSGMLALHEEFMVPAKALNLPARTVTADSRITSDSPMTNINYGT
mouse_5_corrected -----

rat EVS-TVNPQTLOSEHLPDFKLPSTVINGTAVTKSMNPSIASKIEDTINQNPILIIFF--SV
human_5+3_corrected EPSFVVNSQILPPEEPTDFKLSTAIKTAMSKNINPTMSSQIQGITNQHSSTVFPILLGA
mouse_5_corrected -----

rat ARIRDSAQAGRAS--SQSAHPVTGGNMTYGHINTYSSFTSKASTVLQPINPTESYGPQI
human_5+3_corrected TEFQDSQMGGRGHEHFQSRPPITVTRMIKDVNVKMLSSTTNKL--LLESVNTTNSH--QT
mouse_5_corrected -----

rat PITGVSRPSSSDISSHTADPFSSEHPSGSHETASSLPHIPRNNTNPNPLSRHLGREIT
human_5+3_corrected SVREVSEPRNHFYSHTITQILSTSTFPSPDPTAHSQFPPIRNS-TVNIPLFRFRGRORR
mouse_5_corrected -----

rat IWSRGVKNFHRTPVLRHRHRTVPAIKGPANKNVSQVPATEYPMCHTCPSAEGLTVA
human_5+3_corrected IGGRIIISPYRTFVLRHRHYSIFASTTRGSSEKSTTAFSATVLMVTCLSLPRERLITA
mouse_5_corrected -----

rat TAALSVPSSESSALPKTNVGVIAZESTTVKKPILLFKDKQNVDEIIITTTKYSGGES
human_5+3_corrected TAALSFPSAAPITFPKADIARVPSEESTTLVQNPILLLENKP--SVENTPTTIKYFRTEI
mouse_5_corrected -----

rat NHVIPTASMTSAPTSVSLGKSPVDNSGHLMPGTIOTGQDSVETTELPSPLSTP--SIP
human_5+3_corrected SQVPTTGAVMTYAPTSTIMEKTHKVNASYPRVSSSTMEAKRDSVITSSLSGAITKPPMTII
mouse_5_corrected -----

rat TSTKFSKRRTPLEQIFVNNQKKEGMLKNPYQFGLQKNPAKLPKLIAPLLPTGQSSPSDST
human_5+3_corrected AITRFSRRTIPWQONFVNNHNPGRRLRNQHKVSLQKSTAVMLPKTSPALPQROSSPFHFT
mouse_5_corrected -----

rat TLLTSPFPALSTMAATONKGTVEVVGARSLSAGKQ-PYTNSSPVLPTSTIKRSNTLNF
human_5+3_corrected TLISTVMQIPSNLTATAHHTTTKTHNPG-SLPT-KKELPFPPLNPLPSIISKDSSTKSI
mouse_5_corrected -----

rat LSTETPT-VTSPTATASVIMSETQRTSKEAKDQIKG-F-RNNENNANTTPROVSGYSAY
human_5+3_corrected ISTQTALPATTPTFPASVITYETQTERSRAQTIQREQEPQKERTDPNISPDQSSGFTTP
mouse_5_corrected -----

rat SALITADTFLAFSHSPRODGGNVSAYAYESTTS--LLAITELEFKYTQTLGNITTALETT
human_5+3_corrected TAMTP--PALAFTHSPPENTTIGISSTISFHSRTINLTDVIEBLAQASTQTLKSTIASETT
mouse_5_corrected -----

rat LLSKSQESTTVKRAS-DTP-PFLLS8GAPPVETPSPPPFETKGVVTDGKVTSAFQMTSNRV
human_5+3_corrected LSSKSHQSTTTTKASLDTPIPPFLSSSATIMPVPISPFTQRAVTDTRGDSDHFRIMTNTV
mouse_5_corrected -----

rat VTIYESSRHNTDLOQP9AEASPNPEIITGTTDSPNLFSTSVPALRVDPKONSKWKESP
human_5+3_corrected VKLHESSRN--LQMPSSQLEP-----LTSSTENLLHSTMPALITTVKSONSKLATPSP
mouse_5_corrected -----

rat WPEHKYQLESYSETIEKGRPAVMSPELSPLEASTHASHWNTQKHAESVFDKPKGQNP
human_5+3_corrected WAEYQFWHKPYSIDAEKGRPEVSMIATTGLSEATTILVSDWDGQRTKASDFDKKPVQEA
mouse_5_corrected -----

rat -TSKHLPPYVSLPKTLKKPRIIGGKAASFVPAKSDVFLPCEAVGOPLPIIHWTAVSSGX
human_5+3_corrected TSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAFLPCEAVGNELPTIHWTRVS-GL
mouse_5_corrected -----

rat KISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFFVLSVVFYFARILD
human_5+3_corrected DLSRGNQNSRVQVLPNGTLSIQRVSIQDRGQYLCSASKLFGTDHLHVTLVSVVYPPRILE
mouse_5_corrected -----

rat RHVKRITVHFGSTVELKCRVEGMPRTVSWILANQTVVSETAKGSRKVVWTPDGTLLIYN
human_5+3_corrected RRTKEITVHSGSTVELKCRABGRPSPTVTWILANQTVVSESQGSRAVVTVDGTLVLEN
mouse_5_corrected -----

rat LSLYDRGFYKCVASNPGGQDSLLVKIQVITAPPVITEQKROAIVGVLGSSKLKLPCTAKGT
human_5+3_corrected LSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRRQVIVGTWGESLKLKLPCTAKGT
mouse_5_corrected -----

rat PQPSVHWVLYDGTTELKPLQLTHSRFFLYPMGTLYIRSIAPSVRGTYECIATSSSGSERRV
human_5+3_corrected PQPSVYVWVLSDGTVEVKELQFTNSKLFLFNGTLYIRNLASSIRGTYECIATSSSGSERRV
mouse_5_corrected -----

rat VILTVEGETIPRIETASQKTEVNLGEKLLNCSATGDPKPRIIWRLPKAVIDQWHRM
human_5+3_corrected VMLTMEERVTSPIEASQKTEVNFPGDKLLNCSATGEKPKQIMWRLPKAVVDQ-----
mouse_5_corrected -----

rat GSRHVYPNGSLVVGSVTEKDAGDYLCVARNKMGDDLILMHVRLRLTPAKIEQKQYFKQ
human_5+3_corrected GSNHVYPNGSLFIGSVTEKDSGVYLCVARNKMGDDLILMHVSLRLKPAKIDKQYFRKQ
mouse_5_corrected -----

rat VLEGKDFQVDCRAGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTLFHNCTLYFNWV
human_5+3_corrected VLEGKDFQVDCRAGSPVPEISWSLPDGTNINNAMQADDSGHRTRRYTLFNNGTLYFNKV
mouse_5_corrected -----

rat GVAEEGDYTCYAQNTLGKDEMKVHLTVITAIPIRQSYKTMRLRAGETAVIDCEVTGEP
human_5+3_corrected GVAEEGDYTCYAQNTLGKDEMKVHLTVITAIPIRQSNKTKRIKAGETAVIDCEVTGDP
mouse_5_corrected -----

rat KPNVFWLLPSNNVISFSNDPFTTHANRSLSIHKVKPLDSGDYVCVAQNPSGDDTKYKLD
human_5+3_corrected KPKLFWLLPSNDMLSFSDRYTTHANGSLTINKVKLLDSGEYVCVARNPSGDDTKMYKLD
mouse_5_corrected -----

rat IVSKPPLINGLYANKTVIKATAIRHSKXYFDCRADGIPSSQVTWIMPGNIFLPAPYTGSR
human_5+3_corrected VVSKPPLINGLYTNRTVIKATAVIRHSKXHFDCRAEGTPSPEVMWIMPONIFLTAFYTGSR
mouse_5_corrected -----

rat VTVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLMLRRPTFRNPFNEKVIA
human_5+3_corrected ITVHKNGTLEIRNVRLSDSADFTCVARNEGGESVLVVQLEVLMLRRPTFRNPFNEKIVA
mouse_5_corrected -----

rat QAGKPVALNCSVDGNPPPEITWILPDGTQFANRPNSPYIMAGNSLILYKATRNKSGKY
human_5+3_corrected QLGKSTALNCSVDGNPPPEITWILPNGTRFSNGPQSYQYLIASNGSPITISKTREDAGKY
mouse_5_corrected -----

rat RCAARNKVGYIEKLILLEIGQKPVILTYEPGVKSVSGEPLSLHCVSDGIPKPNVKTTP
human_5+3_corrected RCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGBSLSLHCVSDGIPKPNIKWTP
mouse_5_corrected -----

rat GGHVIDRPQVDGKYILHENGTLVIKATTAHDOGNYICRAQNSVGQAVIEVSVMVVAIRPK
human_5+3_corrected SGYVVDRPQINGKYILHENGTLVIKEATAYDRGNYICRAQNSVGHTLITVPVMIVAYPPR
mouse_5_corrected -----

rat IINYLERNMLRRTGEANOLHCVALGIPKPKVTWETPRESLLSKATARKPRSEMLHPQGT
human_5+3_corrected ITNRPPRSIVTETGAAPQLHCVALGVKPEITWEMPDSLLSTASKERTHGSEQLHLOGT
mouse_5_corrected -----

rat LVIQNLQTSDSGVYKCRQNLSTDYATTYIQV
human_5+3_corrected LVIQNPQTSDSGIYKCTAKNPLGSDYAATTYIQV
mouse_5_corrected -----

(rat: SEQ ID NO:10)

(human_5+3_corrected: SEQ ID NO:11)

(mouse_5_corrected: SEQ ID NO:12)

[illegible]

[illegible]

rat	SRIHVYPNGSLVVGSVTERDAGDYL CVARNKNGDOLVIMHVRLRLTPAKIBQKQYFKQV
human_5+3_corrected	SWIHVYPNGSLFIGSVTERKSGVYLCVARNKNGDOLIMHVSLRLPKAKIDBKQYFRQV * *****,;*****;* *****,;**** ***,****;****;***
rat	LHGKDFQVDCASGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGLTYFNNVG
human_5+3_corrected	LHGKDFQVDCASGSPVPEISWSLPDGTMINNAMQADDSGERTRRYTLFHNGLTYFNVG *****;*****;***. *****,;****;*****;***
rat	MAEGDYICSAQNTLGKDEMKVHLTVLTAIPRIQSYKTIMRLRAGSTAVLDCEVTGEPK
human_5+3_corrected	VAEGDYTCYAQNTLGKDEMKVHLTVITAAPRIQSNKTNKRIKAGDTAVLDCEVTGDEK ;***** * *****,;*** ***** *. *;***;*****;***
rat	PNVFWLLPSNNVISFSNDRFTFHANRTLSIHKVKPLDSGDYVCVAQNSGDDTKTYKLDI
human_5+3_corrected	PKI FWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPSGDDTKMYKLDV *;;*****;:**** *;***** *;*** *****;***** *****
rat	VSKPPLINGLYANKTVIKATAIRHSKXYFDCRADGIFSSQVTWIMPONI FLPAFYGSRV
human_5+3_corrected	VSKPPLINGLYNRTVIKATAVRHSKXHFDCRAEGTSPSEVMWIMPONI FLTAPYGSRI *****;***;*****;*****;**** *;*** *****;****;***
rat	TVHNGTLEMRNIRLSDSADFTCVVRSEGGEVVLVQLEVLEMLRRPTFRNPFNEKIVAQ
human_5+3_corrected	TVHNGTLEIRNVRLESDSADFTCVARNEGGESVVLVQLEVLEMLRRPTFRNPFNEKIVAQ *** *****,;***** ***** *. *****;*****;***
rat	AGKPVALNCSVDGNPPPEITWILPDGTQFANRPHNSFYLMAGNSLILYKATRNKSGKYR
human_5+3_corrected	LKSTALNCSVDGNPPPEI IWILPNGTREFSNGEQSYQYLIASNGSPIISKYTRDAGKYR ** ..***** ***** *****,;*** *;*. ***,;****;*. *;***;****
rat	CAARNKVGYTEKLILLEIGQKPVILTYEPMVKSVSSEPLSLHCVSDGIPKPNVKWTPG
human_5+3_corrected	CAARNKVGYTEKLIVILEIGQKPVILTYAEGTVKGISGESLSLHCVSDGIPKPNIKWTMP *****;***** ***** ** *,;*** *****;*** *
rat	GEVIDRQVDGKYLHENGTLVIKATTAHDQNYICRAQNSVGQAVLSVSVNVVAYPPRI
human_5+3_corrected	GYVDRPQINGKYLLHDNGTLVKEATAYDRGNYICKAQN SVGHTLITVPVMIVAYPPRI *;*;*****;*****;***** *;***;*****;*****;***;***;*****
rat	INYLFRNMLRRTGEAMQLHCVLGIPIPKVYTWETPRHSLLSKATARKPRSEMLHQQTL
human_5+3_corrected	TNRPPASIVTRTGAFLHCVLGVKPEITWEMPDHSLSTASKERTHGEQHLHQQTL * **,:* ** *;*****;***;*** * *****,*! ..* ** ** *
rat	VIQNLQTSDSGVYACRAQNLGTDYATYIQVL
human_5+3_corrected	VIQNPQTSDSGIYKCTAKNPLGSDYATYIQVI **** *****;*** *;* *****,;*****;

(rat: SEQ ID NO:13)

(human_5+3_corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLSLQVLKMSYNKVQIIE
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQIF
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESYLHGNPWTCDCHLKWLSEWMQGNP
(SEQ ID NO: 15)

Figure 18

MKVKGGRGITCLLVSFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSIPDSIPPNVE
RINLGYNSLVRLMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
QIFKISFIKFLYLSDNFLTSLPQEMSYMPDLSLLYLHGNPWTCDCHLRWLSDWIQPDVI
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFSTFL
VCNIDYGHIPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
ADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEPRVVKHKWTMISRDNNTK
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
ADSFDTG VYHCISSNYDDADILTYRITVVEPLVEAYQENGIIHTVFIGETLDLPCHSTGIP
DASISWVIPGNVLYQSSRDKKVLNNGTLRLQVTPKDQGYRCAANPSGVDFLIFQV
SVKMKGQRPLEHDGETEGSGLDESNPJIAHLKEPPGAQLRTSALMEAEVGKHTSSTSKRH
NYRELTQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALLEKAKKNAMPDKRENTTV
SPPPVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRITSDSPMTNIN
YGTEFSPVVNSQILPPEEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLG
ATEFQSDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV
SEPRHNHFYSHTTQILSTSTFSPDPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGRGRIS
PYRTPVLRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
SIPMEKTHKVNASYPVSVSTNEAKRDSVITSSLSGAIKPPMTIIAITRFSRRKIPWQQNFV
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNTLTT
AHHTTTKTHNPGSLPTKKELPFPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
TQTERSRAQTIQREQEPQKKNRTDPNISPQSSGFTTPTAMTPPALAFTHSPPENTTGISST
ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETTLSSKSHQSTTTRKASLDTPIPPFLSSSAT
LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLQMPSSQLEPLTSSTSNLL
HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGGKPEVSMLATTGLSEATTLV
SDWDGQKNTKKSDFDKKPVQEATTSKLLPFDSLRYIFEKPRIVGGKAASFIPANSDAF
LPCEAVGNPLPTIHWTRVSGDLDSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCSASNL
FGTDHLHVTLSVVSYPRIERRTKEITVHSGSTVELKCRAGEGRSPPTVTWILANQTVVSE
SSQGSRQAVVTVDGTLVLHNLISIYDRGFYKCVASNPGGQDSSLVKIQVIAAPPVILEQRR
QVIVGTWGESLKLPTAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSSTGSERRVVMLTMEERVTSPIEAASQKRTEVNFGDKLLLNCSATGEP
 KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCASGSPVPEISWSLPDGTMINNAMQADDSGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLLTVITAAPRIRQSNKTN
 KRIKAGDTAAVLDC EVTGD PKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPSGDDTKMYKLDVVSKPPLINGLYTNRTVIKATAVRHSHKHFDCRAEGTP
 SPEVMWIMPDNIFLTAPYYGSRTVHKNGTLEIRNVRLSADFICVARNEGGEVLVVQLE
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSVDGNPPPEIHWLPNGTRFSNGPQSY
 QYLIASNGSFIISKTTREDAGKYRCAARNKVG YIEKL VILEIGQKPVILTYAPGTVKGISGE
 SLSLHCVS DGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATAYDRGNYICKA
 QNSVGHTLITVPVMIVA YPPRITNRPPRSIVTRTGAA FQLHCVALGVPKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGT LVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
 NO: 16)

Figure 19

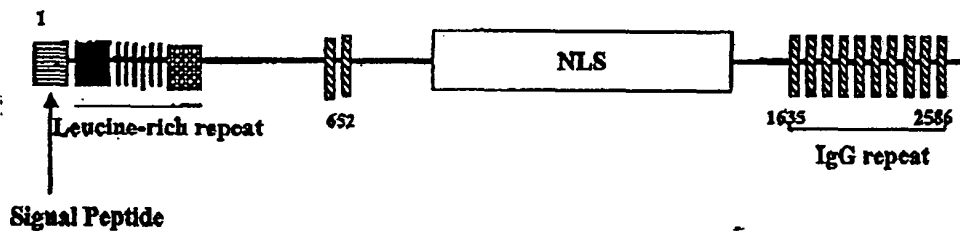


Figure 20

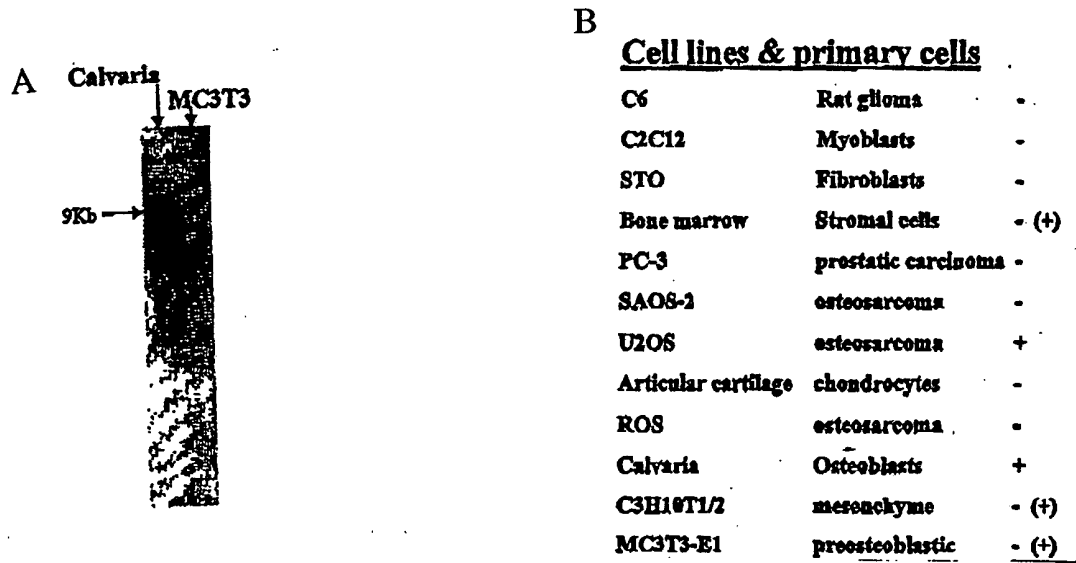


Figure 21

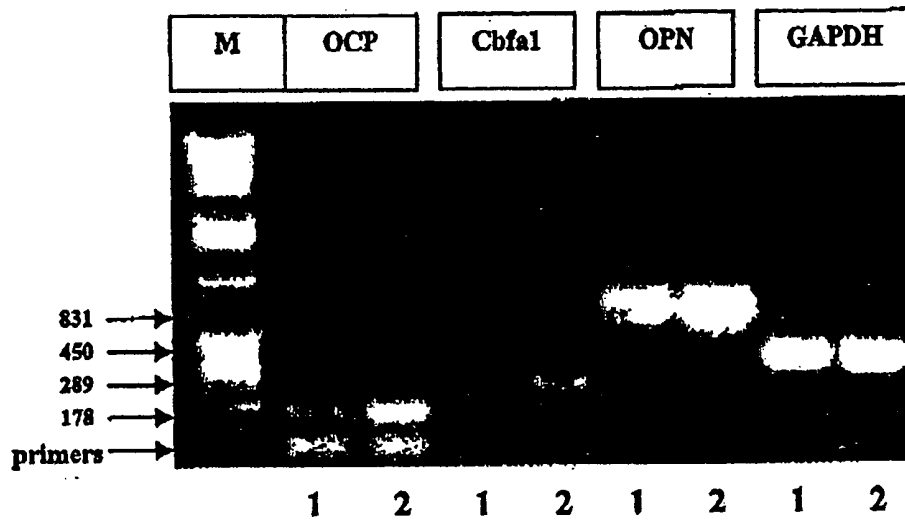


Figure 22

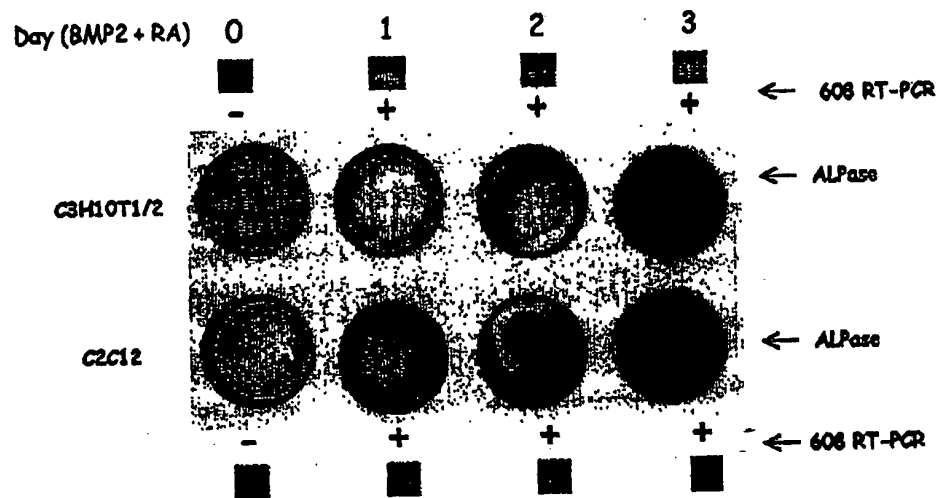


Figure 23

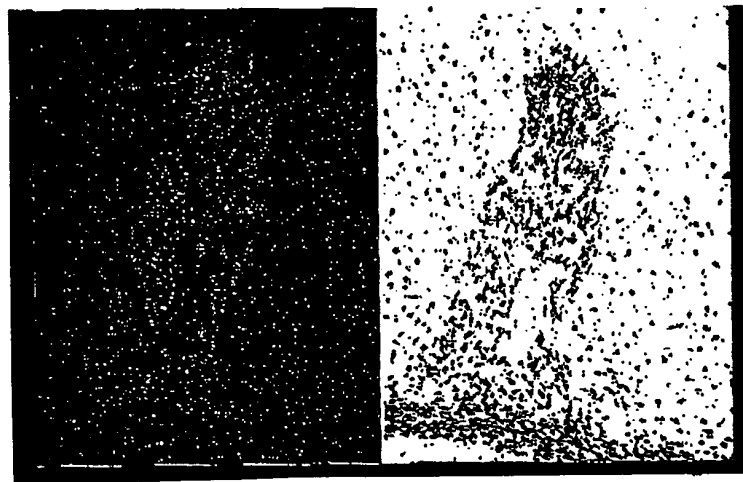


Figure 24

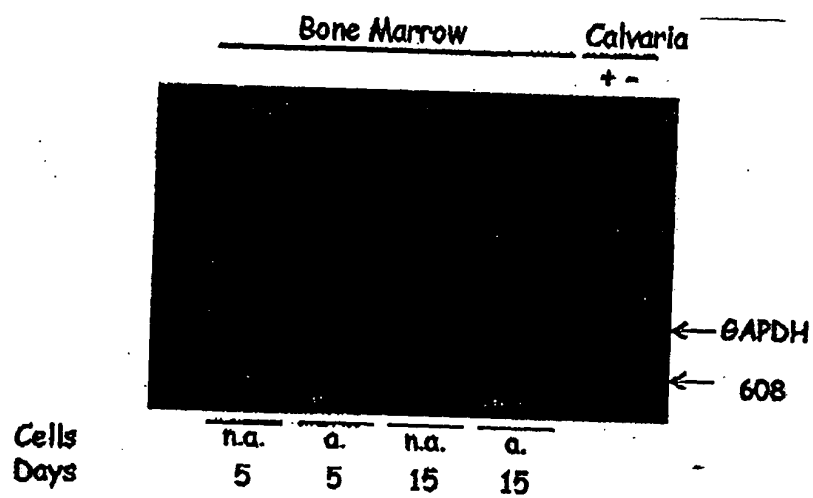


Figure 25

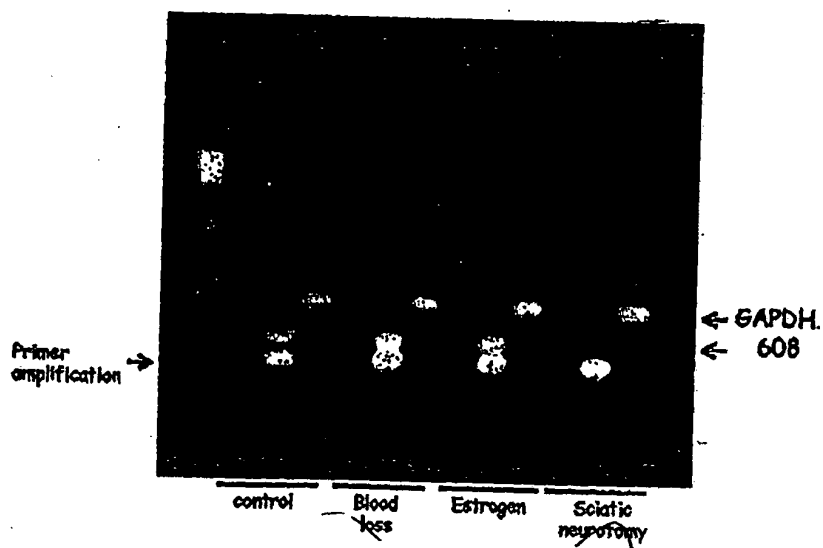


Figure 26



Figure 26

Figure 27

A

B

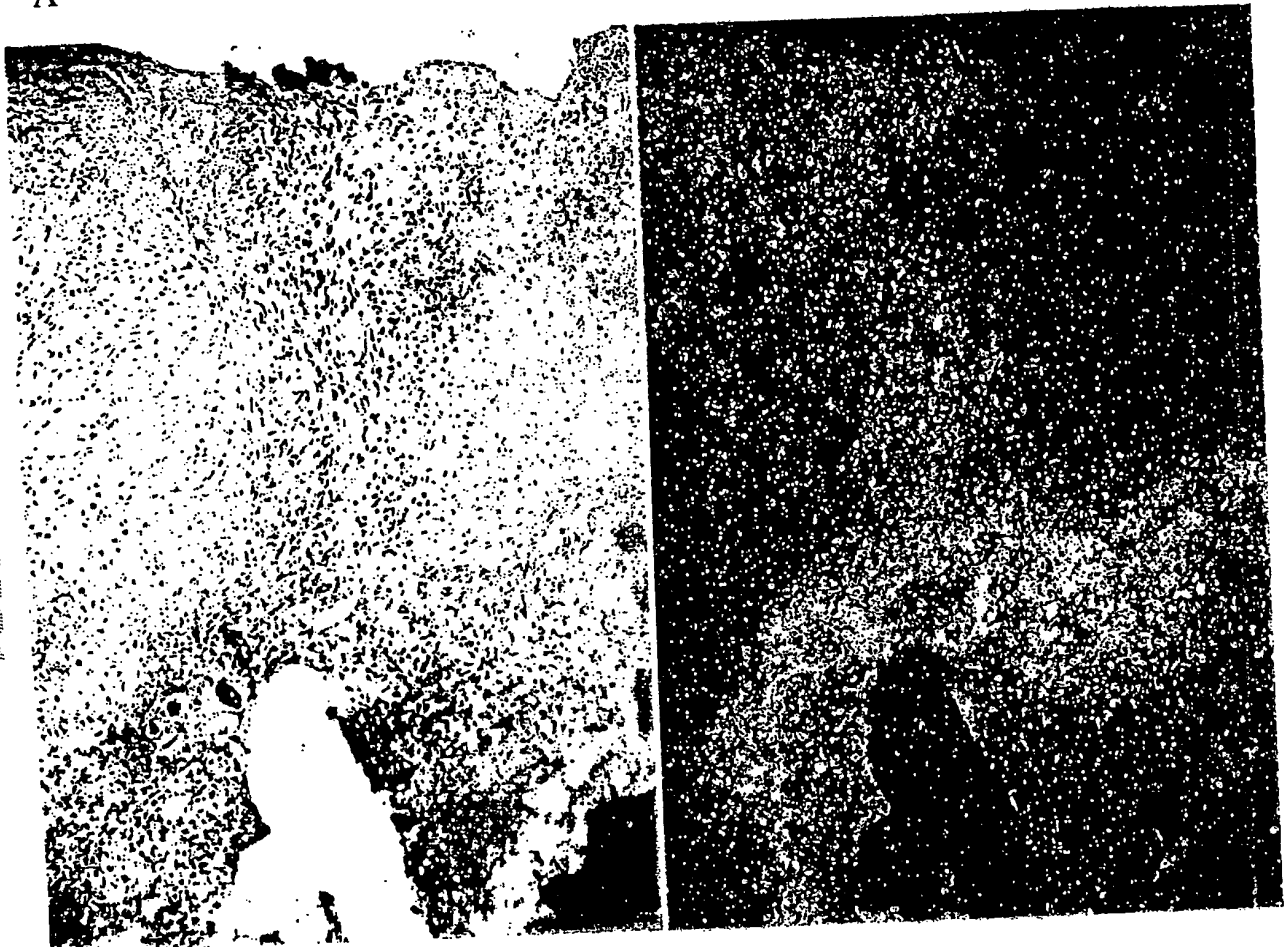


Figure 28



Figure 29

A



B



Figure 30

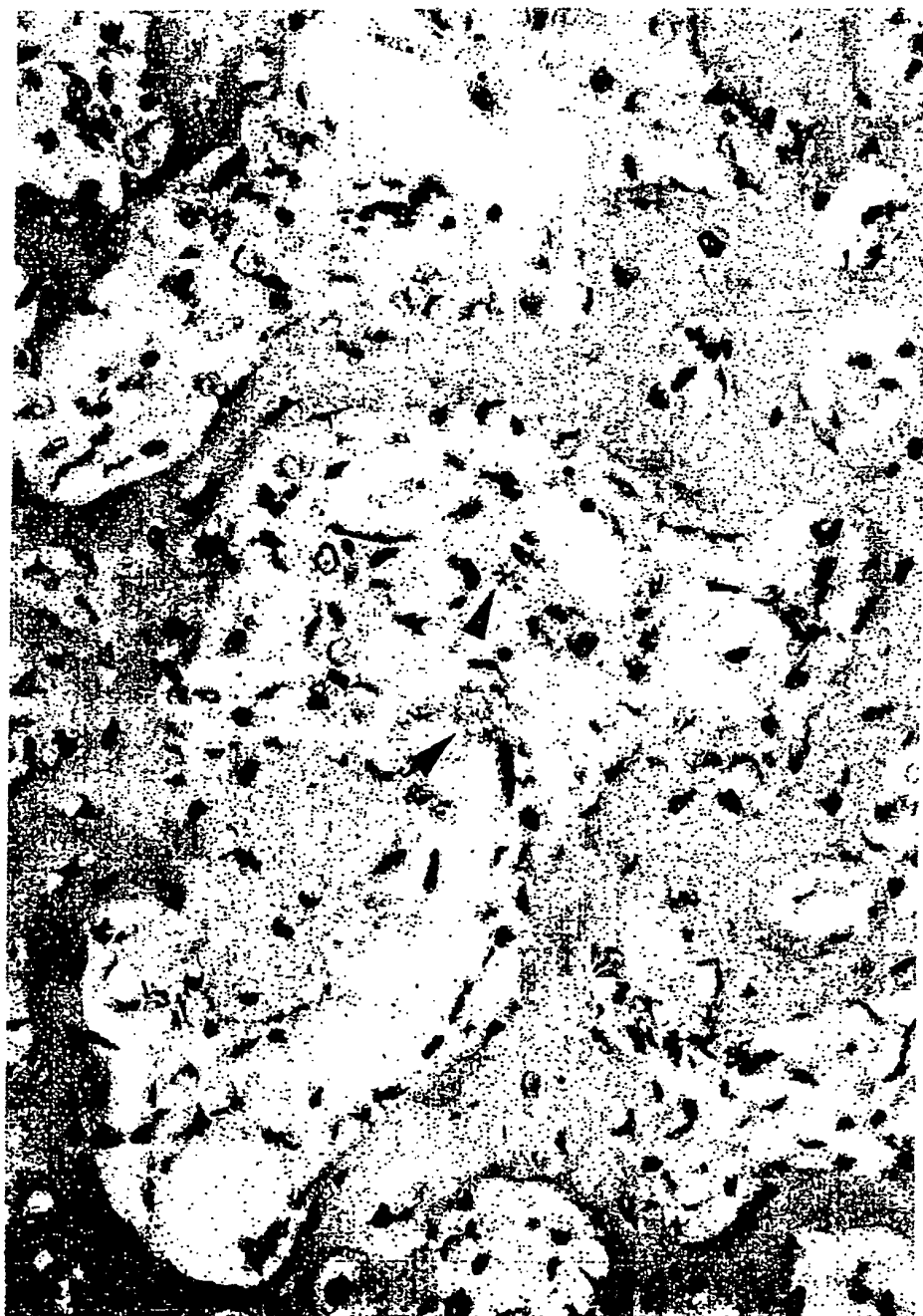
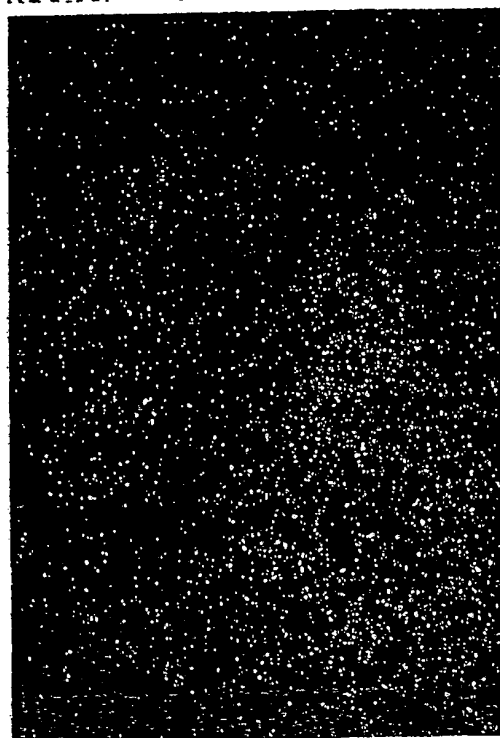


Figure 31



1000X 1000X 1000X

Species	Age	Sex	Weight (g)	Length (mm)	Wing (mm)	Tail (mm)	Culmen (mm)	Gape (mm)	Bill (mm)	Foot (mm)	Middle toe (mm)	Claw (mm)	Toe + claw (mm)	Weight (g)	Length (mm)	Wing (mm)	Tail (mm)	Culmen (mm)	Gape (mm)	Bill (mm)	Foot (mm)	Middle toe (mm)	Claw (mm)	Toe + claw (mm)
1	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
2	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
3	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
4	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
5	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
6	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
7	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
8	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
9	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
10	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
11	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
12	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
13	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
14	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
15	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
16	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
17	1	♂	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
18	1	♀	10.5	110	35	25	10	10	10	10	10	10	10	10	110	35	25	10	10	10	10	10	10	10
19	1	♂	10.5	110	35	25	10	10	1															



B

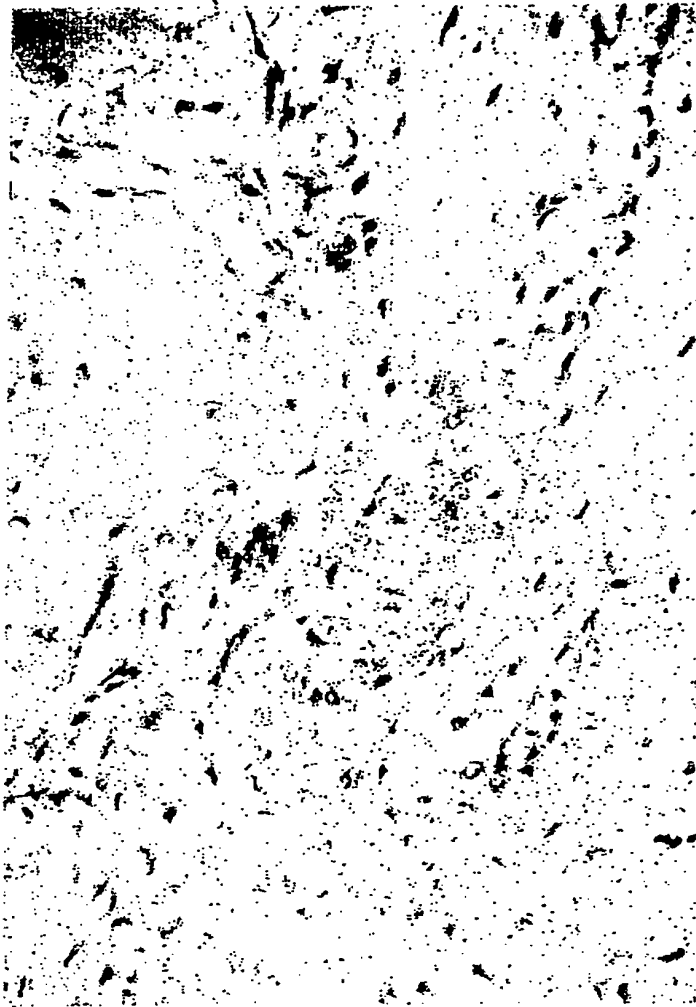


Figure 34

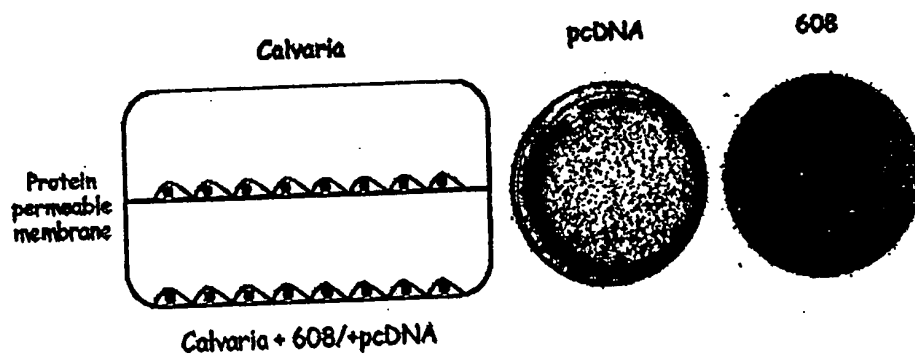


Figure 35

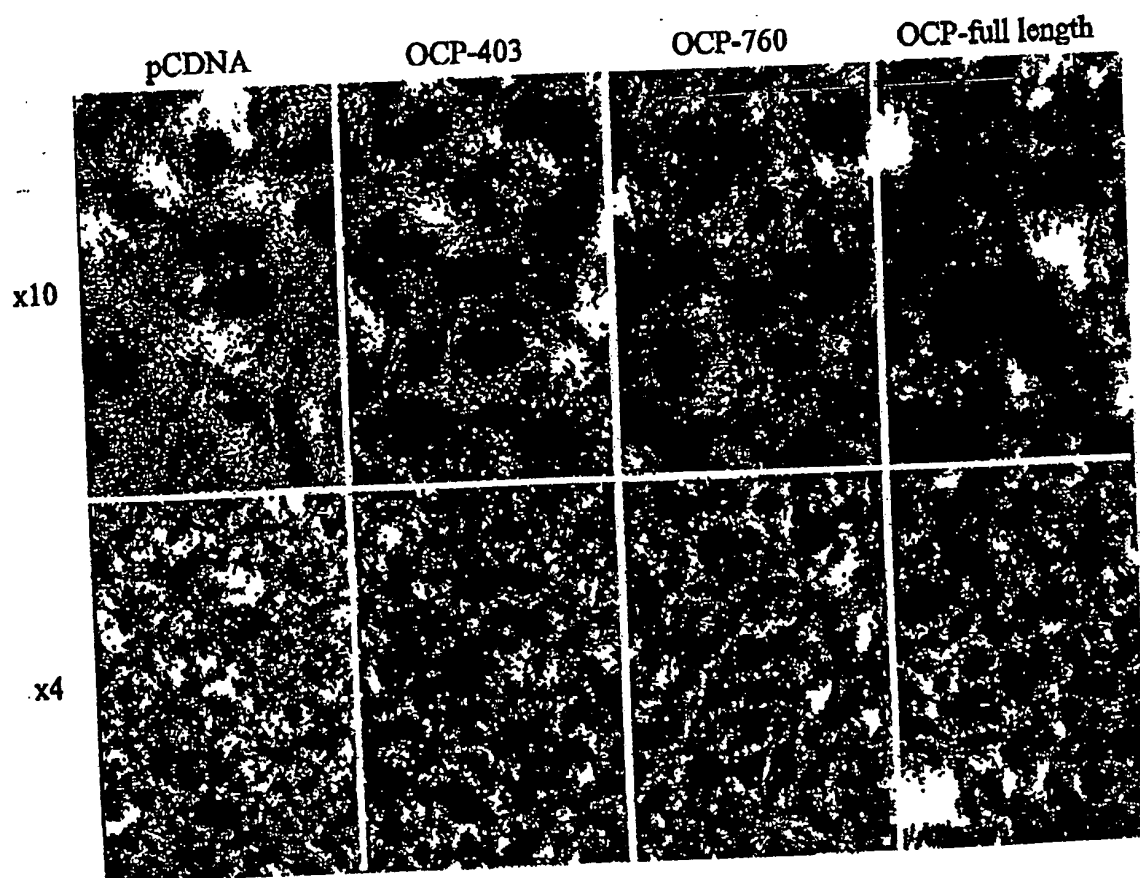


Figure 36

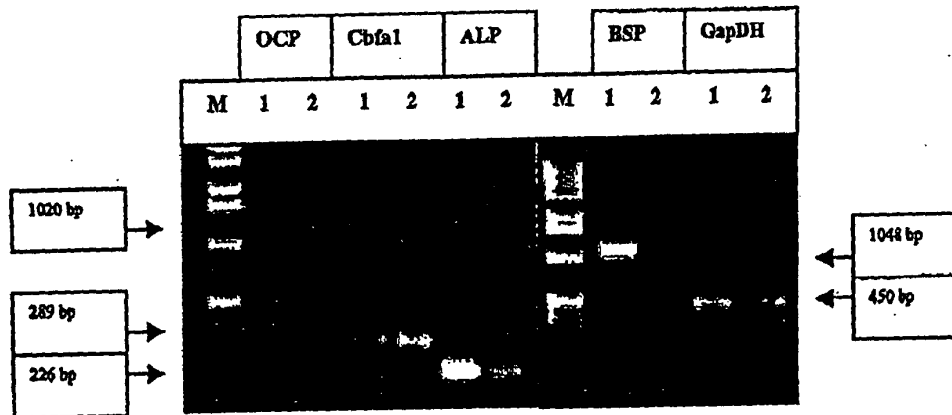


Figure 37

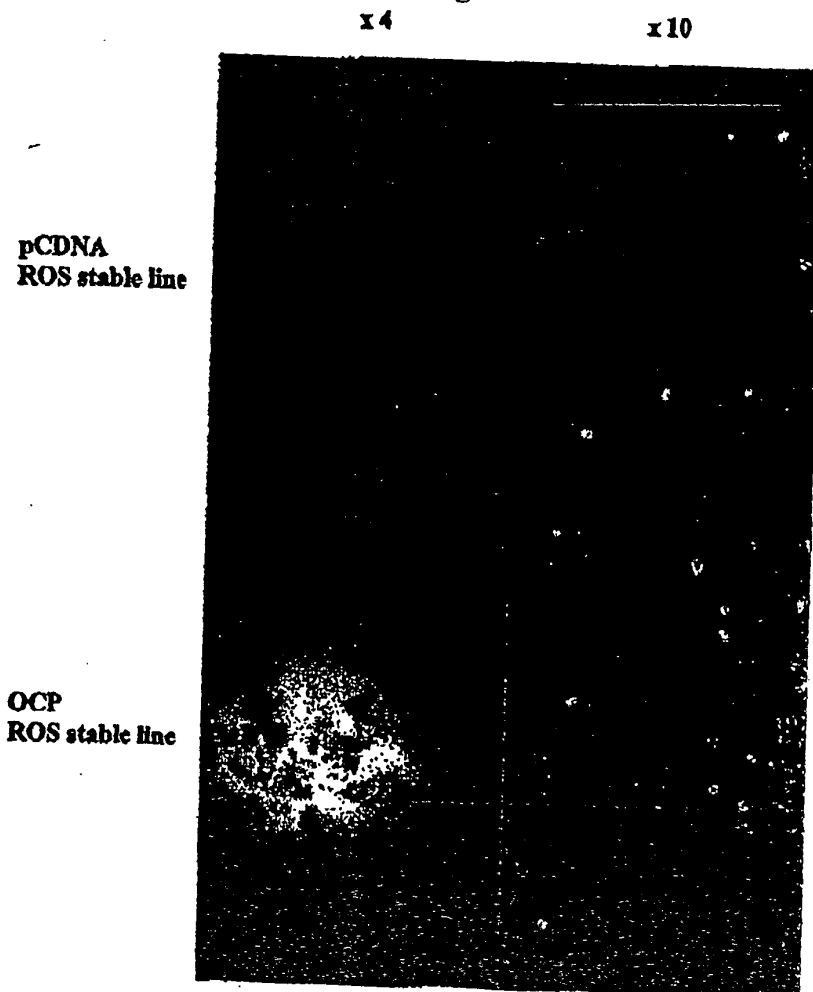


Figure 38

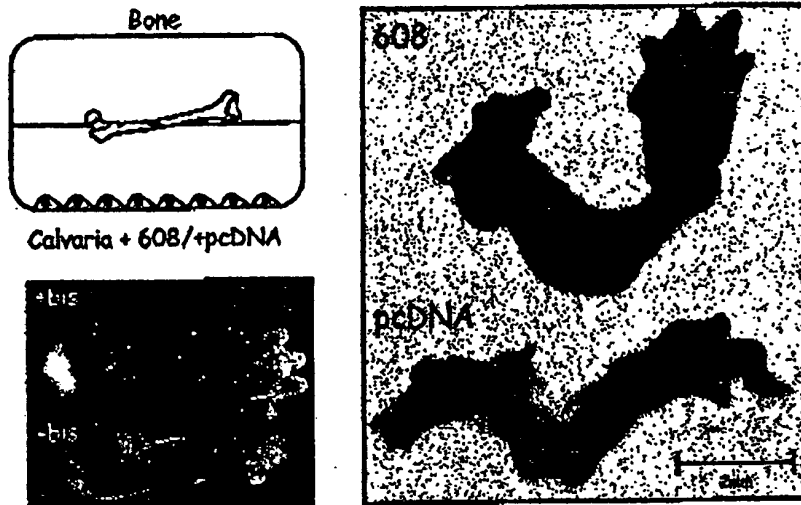


Figure 39

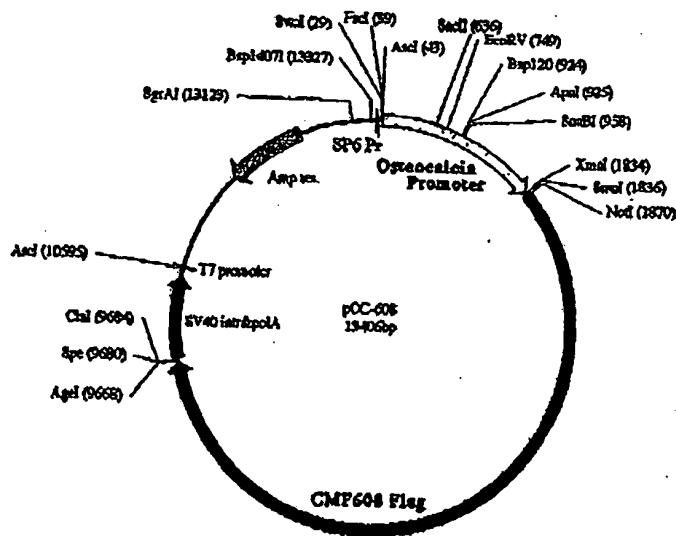


Figure 40

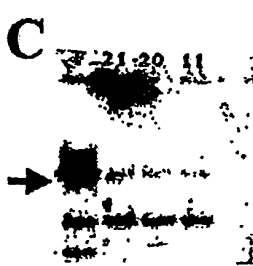
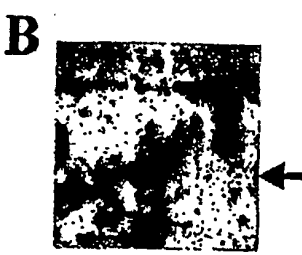
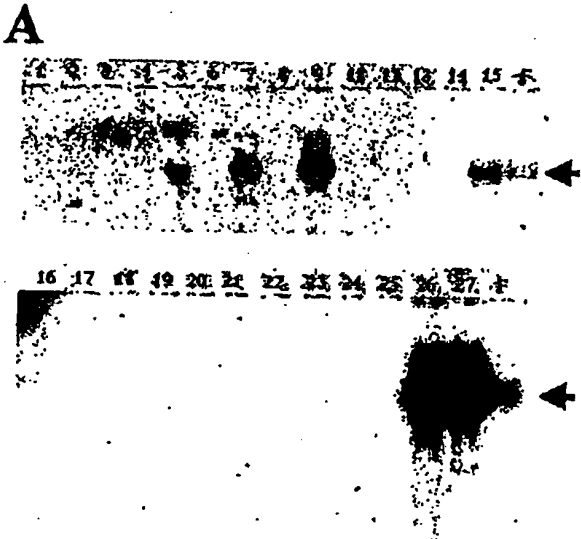


Figure 41

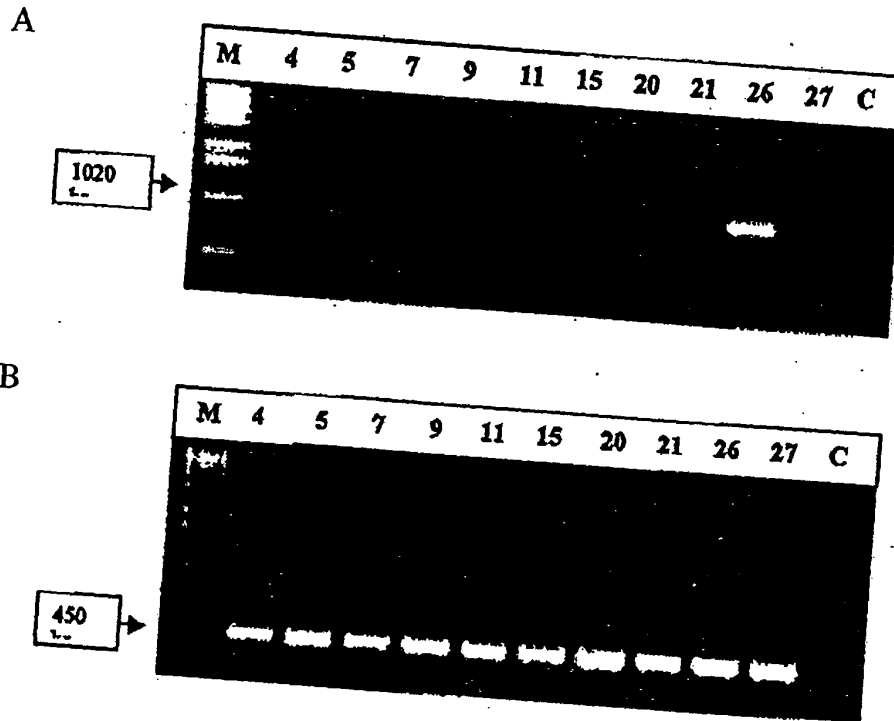
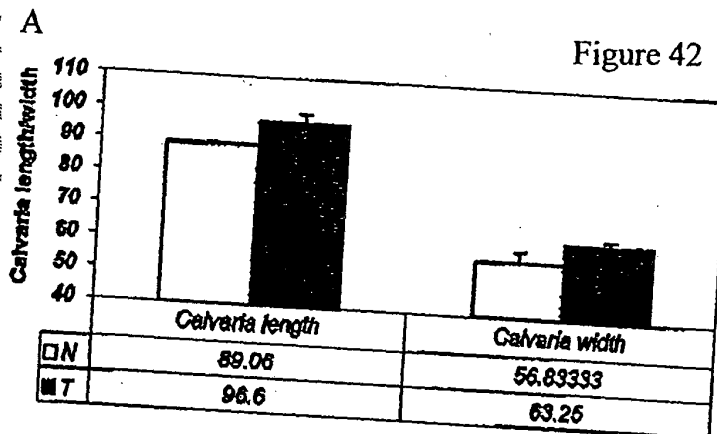


Figure 42



B

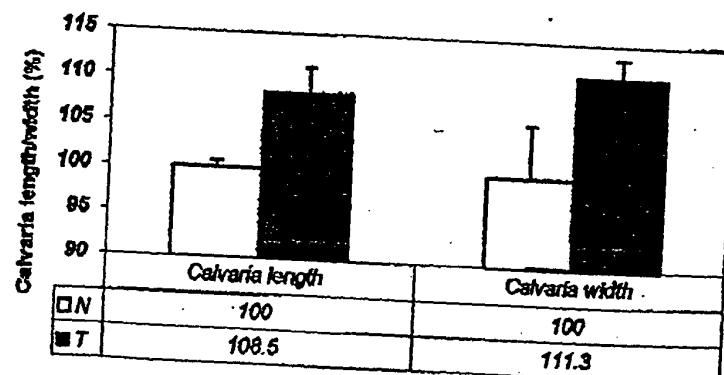


Figure 43

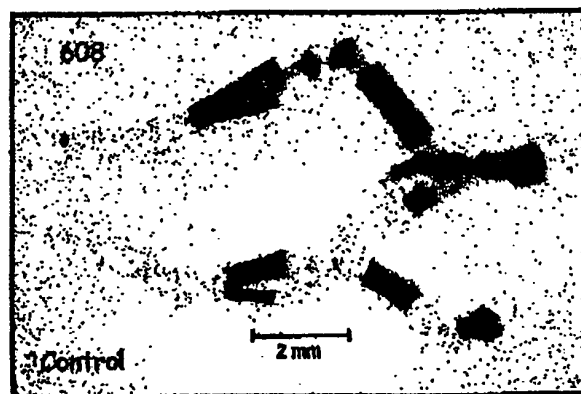


Figure 44

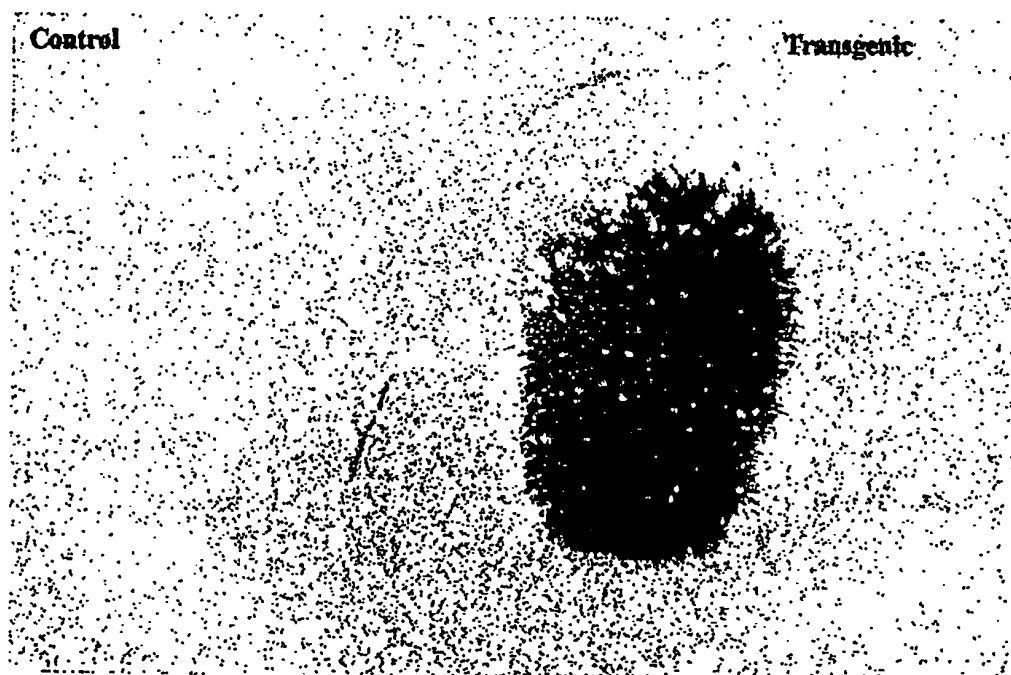
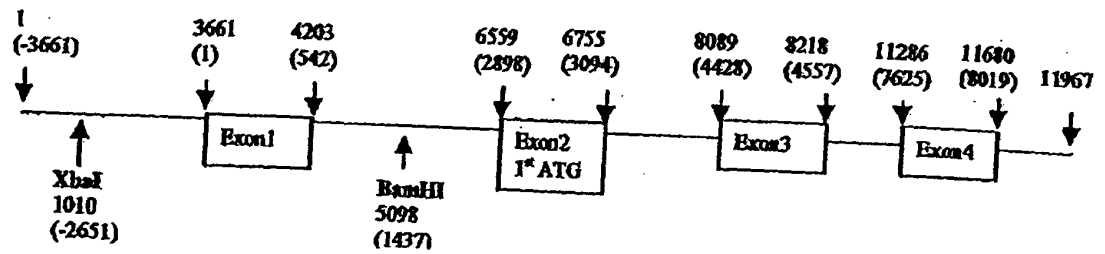


Figure 45



Clone 14C10 /15E11

T7 orientation in pKS

Figure 46

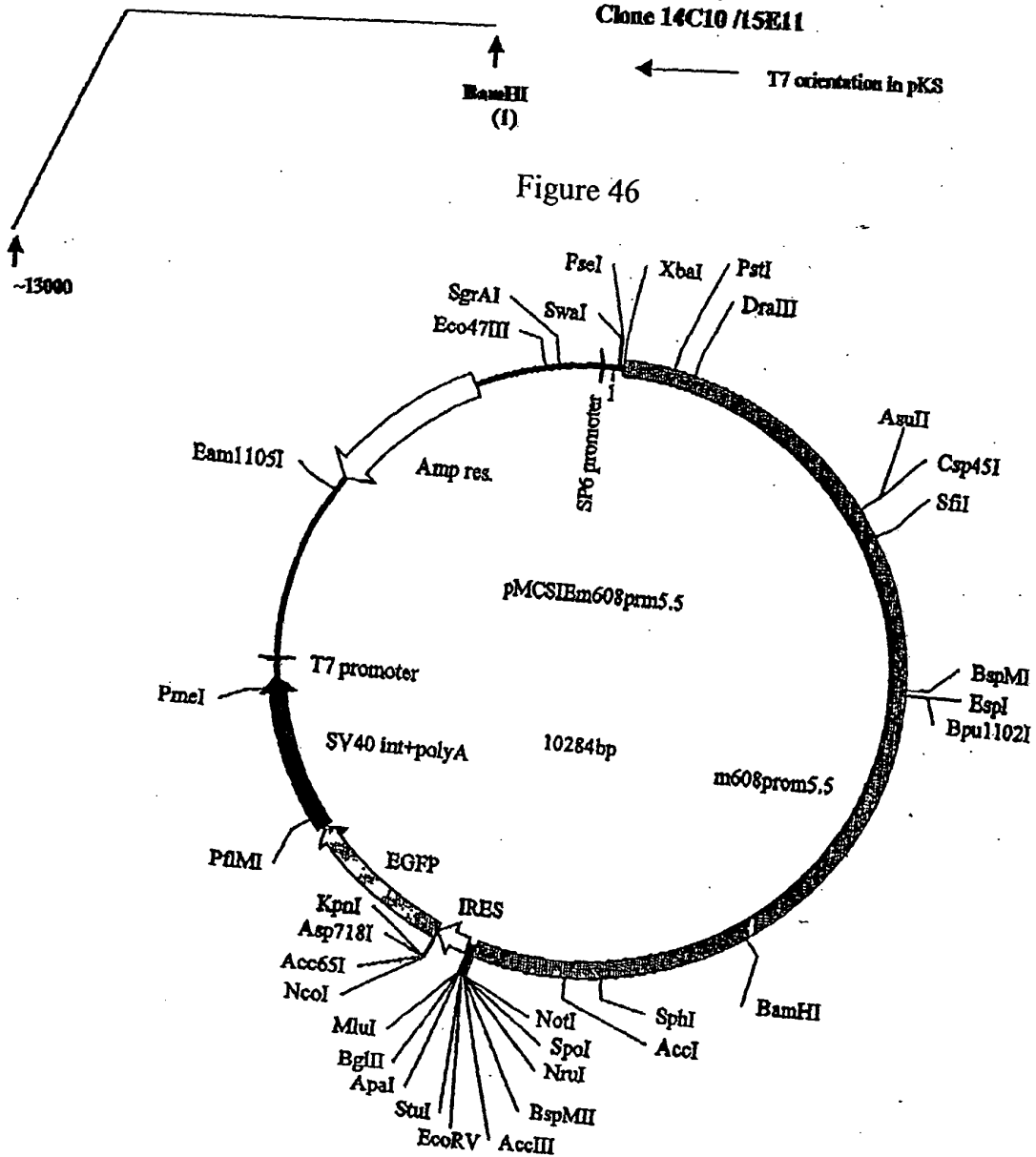


Figure 47

↓ (XbaI)

TCTAGAGTAAATGATCCTGAGTAGCGATCCTGGGAAAAATACGTA CTCTAACACACTGC
 AATCATCTCTCTGTGGTTTGTCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATA
 ACCTACCGAATACAGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCT
 GTAGCTGCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCACATGCCGACAT
 CTTAAGAATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCGTGCCACACAGGTCTCCAG
 GGTIAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTACCTTGATCATTAACTAGTAGAA
 ATCACTTGGTGATTAGCAACTGGATATGGAATATTACTAATTTGTACCCAAAGCCAGGCCAC
 CTCAGCTTTGGCAGCTCCATTCACTCTGTGGAGCCAGTCACGTGGGTTTGAATCAACTGT
 ACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTTCACAAGCTTTTTTA
 GAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATG
 TGTTCATGTATGCTCATGGATGCTATTATAATGTGGAAATAAACTGACTTTAAAAA
 AGCTTATGATACTTGTCACAGAGTAAATCTTCCATAAAATATCATCTGCATTTATAAATAT
 TTTCAATAOCATCAATTA AAAACCTTTAGAAAATTTGTTAACACAAAGATCCCTAGGCC
 CTGCCCTAGGATGGTCTGTATGGTGGGCTGAGAGATGGAGCTTAAGAACTTACTTGCTCC
 AGGAGCACATCTTCAGAACATCTGCCTCAAAAACATTTATCCCAAATGCTCATCAAAGGCTC
 ACTCACATGTGCTTCAACCACAGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGT
 GGAAGCCTGCTAGAGGAACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTC
 ATAACCTCACCAACAACAAACAACGACAACAAACCCATAAAATATCAAGGCAAAAT
 GAACTAAGCCATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTITTTCTTAAT
 AGTCAAAATATGCAGAAATATAGCAAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATA
 CTTGTAGATTAGAGGACACAAAGCAACAACCTACAGAGTGAAAGTAATCCAGAGACTTAATG
 TATAATATGAGGACTGTATTTAATAATTCTATTAAAGATACACAGCAAAAGAGTGATCTT
 ACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGTITTTATCTTCATG
 TAGCTGATAATTTCATATTGTACACCTCAAACATAGATAACCAACAAAGAGGAAGAGGAT
 AGGTGCTCTCCAGGGCGGAAGAGTACATTCCGAAAGTCAGACACCATTGTGTAGATGTA
 CCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCTATAGG
 TGGAAACAACATTATGAGCTAACCCAGTACCCCGAGCTCTTGACTCTAGCTGCATATATATC
 AAAAGATGGCCTAATCGGCCATCACTGGAAAGAGAGGCCATTGGACTTGCAAACTTTAT
 ATGCCCCAGTACAGGGGAATACAGGGGCCAAAAAGGGGGAGTGGGTGGGCAGGGGAGTG
 GGGGTGGGTGGATATGGGGGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATAC
 CTAATAAAAAATGGAAAAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG
 TTCCTGCTATCTCAACTTGCAAAACAGAGGGGAGTCACAAACCCAGGACAAACACAGT
 GATTGAAGCGTCTTTGAATGTTATTGCTGTTGTTGTTTACCACCATCATTAGCATATATTCAT
 TGTGAAAAAATTACGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTGAGC
 ATATTTGGCACCACTACAGCCCGAGCCCTTTGCCCCGCCCCAACACACACACACACAC
 CACCTTTACCTTCTCCTGGG
 CATCATCTGCTCACTACCCACCCAAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTCTC
 CTATTTTATGCTAGGTTCCCTCTCCCTGTTAGGAGATGGGAGAGGTCACGAAAGGAA
 AGAATTTGTAGCCCTGAGCCAGCCCGGCCACAGAGCCTGCCACCAGACAGGAAAGGCC
 CAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGGGTGGCGCTGGTGGT
 CCGGGGTGATGGCCCGCCCATTCOCAGAAGCCGAGGCTATAGCTGCGTCACCTGCCCG
 CCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGACCCCAAGGCGGTGCAATTCCA
 TTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCCTCGGGAAAGTGAAGGGACCCAGAG

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTTCAGGGCGAGACAGAGGGTGGGATGGGCGAGAAG
 GGTCCAGGAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAA
 CATCGCTTCAGTGAAGCCAGGCAAAAGGAGCGGGAAAGGATTATGTAGCCTGGGA
 CGCTTTCATAAACTGATGACGTGTTTGTGCAAAAGCAAGCAATTGAGGAGAAACGGCT
 GGGAGTCCGAAAGAGGAGTGATCGATTAGTACTTGAAGTTAGGTGAGTTGAGAAC
 TAACTAACCTATACTATTGAGGGAGAAAGGAAGAGCATTCCAGCAGCAGCAGCAGCA
 GCAATCAGATAAAGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAATAACA
 GAAGCGCTCCAGTTCTCTGAAGAGTCAGTCCCGCAGCTAGTGAAGACTAAGCCTACTAA
 GCCTTTTGCTCCCGTTGGAAGCAAAAGAGTCTCTCAATCAGGTGAAGGCTCTCTCAGA

AGATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAAGCAAGACAGAAGAGCTCAGG
TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTAACGGGAACATCATGTGTA
TTCAATTTACGGCTCTCTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTCTCTGTTTCC
CAGCTACCTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC
AOCGACAGAGTATTCAAAAATATTCAAGTATGTAACAGTGACAGTGTAGGCATAGAAGTA
ATCATTAGTAAATCTTAATTTGGGTAAACTCATTATAACAGCTCCAGGTTGGGAGGGAT
CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAAATCTTC
CTTGCCCATGCTCCCCATCACTGTGTCAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAG
AGCAGACAGTGTATGCCTGCAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAAGTTTACT
GTCTCCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGC
AGACATGCTAGAAAAGAAATTTGACCTTTTCTAGTGGGTATTACAGCTGTAAAAGTAT
TTTGGAAGGTAAAGCCAAATAAATAAAACACATATTAATAATACAATGTTACAAAAT
GATCATATAAAGAAGTACATTCTATAATGCAATGTGAAAAATATATATAATTTTATCTAT
TACTGGTGCAAAGTTTTCTAAATGACATGTACTATTTTATATTTATAAAAAATTTTT
AAAATGTATATAAAGGTGTAAGGGCTCTTGGTCAACAAGAGAGTTAAATTTACAACT

↓ (BamHI)

TTAATTTGTCGGATAACATTATTATGATCTCTAATGACAGGGATCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAAAGCCATTTAGTGATAAAGTCCAATG
GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCTGAGTCTATGCAAGTTTACA
CTTAGGGAATAAAGAGTTCTCTCTCTCCAGATTTCACATGCAATTTGTTGTCATCTTATC
TTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
ATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCTTTATTGGATAA
TTGGGGGATGTTGTTAATTTTTCAGTTCTTATGGAATTCAAAAAACAAAAACAAAC
AAACAAACAAAAACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
AGAAATCTGCACAGAATTTATTGTACATTGTTCAATTATTCACGACAGCCAAGAAATGTGGA
ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAAATGTGGAATGTGTACAAC
AGAGTCCCATGTGGCCATAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
GGAAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTT
CTCTGACAAACAGAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATT
TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACA
AGCTAGCTGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
GGCACACGTTCAAACCAGGGTGAATCCAGCACAGAGAAAGGGGAAAGTAGACACAAAGT
TTGCCCACTAACCAAGAAGCCATTTCAGTTGCTGCTGCTGGGAGGGGCGTTCCAGTTT
CTCCAGTCTGACACTGTGTATAACAACAGTTGACAATACAAAGTTGGCATGATGGATGG
TTTTTGTGCTATTTTTCATTTTCTTACTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTG
GCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAAGGAACATGAAATTTGGGTGGTAA
AGCTGGAAACGATCTGAAAGAAAGTTGGGAAAGAGAAAAATTTGTATGGAGCATATTTAA
ACAAACAAACAAACAAACAAAGGTTCAATTTGCCACAAAAAGGTGTGAATTAATTAAC
CAGTTACGACTCTTAAGAAAAATTTCCCAATTATCCAGAGTTGCTATGTATGCTGTGC
TAGGACTTTGCTTGAACGGCCATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial

NotI site

(SEQ ID NO:17)

Figure 48

CACAGACGCTTCTCTCTCTAACTCTCTCTCCCCATCTTGTTGCTTCATCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCTGTGTCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAACTGCTCTGTGAGAGTCAAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGTATCTCCATTCTTCTCTGACCACATCCCCAAGTTATCACAAGGAGTTT
TTCTCCAAACCTTTCTCTCCAGCAAAACCCCTTCAGCTCCTTGGGTACTTTCTCTTAGCCCT
TCATGGAAGAACCTGTGCTCCATCCCAATGGATGGCTGTGAGCATCCACTCTGTATAGAAT
CTTGGTCAGTGCAGTCTTTTGATATCCCTCAAGAAACACTGGGTCTGAAAAATTTAAACCCAAAG
AACTGTTTTTTGTTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTTGTCTCCTCTCTTTCAGTACCAACCTGCTTTGTCTGCTTTCTCAAAGAATCAAATAAA
ACCAAAAGTGATATAGATTCAATGGCAATCTCTCTAACTGTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAAAGCCTGGAAGCACTCAGAGGCAGTTTACAGAGAACTTAGCCTC
CTGGAACCTTGGCATTCCCATAGCTAGATAAGATGCCCGAGATTGTGCTGTAGATATTGTGTGTG
GGTCTTGTCATGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAAATCTCAGGGCGAGCT
TAGCAAAAGTATACCTAGAATCTTCATGACAGTCAGGTATTGCAAACTACATTGCATATTA
GAAGAAAGTTTGGTAAATTTCTTCTGACAAATGGAGATTTCCTACAGATAAACTTAAAAAGAAC
AGCTAAAGTCACACTCATATGCAAGAATTTACCAAGGCTTAGGAAAGGGGGGGGGGACT
GCTTTATTCTAGTAAAGGTCTGACAGCAAGACCCCTGGTGTAGCTTTTCAAGGTTTCA
AAAGGTGTAGCATAAATTTGTGACTAGAGTGTGAAATCTTTACCTGTCAATTAGCTGACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCTGTCTCAGTCTTTGAAGGCAT
TCCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCTGCTGTGACACCCCTACTCCTTTGT
TTTCTGTATTATATAAGCCCTGGTGTTCCTTTTGTGAAAAATTTACATCCAGATACAGCACTC
CCTTTGTGTCTGTGTCCTTTTGTGACTTTCTGGCCAACTCCATGCCCCACTGCCAGAACCCCTA
GCTTTTTCCACAGATTGAGGGAGGCCGACTGAGCCTGGTTCATGGCATCTAAACCACTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTTGTACACCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTAAATCCCATCTCTTCTCGTTTGTCTAATATAGCAAAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGCTTCCACTCAGAGGGCAGAGCCAGTCTGATCTGACTTTCTGTATGTG
TGTCCATTGTTTCTCTCAGTCTCTGTGTCCTTATAGGAATCTTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGGCTTTCAACGTTTTGCTGT
TAAACAAGATATATGGTGGCAGTTTATAATCATATAAAGCTTTGAAAATAATATATAACTC
CATTCTCATTTCTGCTTCCACGCTGAGCAGTCTCTGTTTCCCCAGGGACCACAAAGAGCATTTGA
AAAGTAGTGATTTATGACCTGCTTTGTCTGTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACCTTGGGGGTGACCTGTATCTGTGTTCTGGGTCTGGTCTCACTATA
TTTGTCTCCAGAATAAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGGAGTCCCAOCATTCTACTTCTGTAACTGTATATGGTAGAAAAAGTTAATTT
ATGTGATTCTTCCAACTCCAAATATTTCAAATCTTTTAGCCCCCTCAGCCTGGGATTTCTTG
ACTAAGTCTATTGATTGGAAGATCTCAGTGGTTAGGATTTGCAGTCATGATGTTTCATACG
TCAGGCTAAGCTGAAAAATATGACAAATGAAATGAAATGTCTATGCTCGGGGAATGTG
AGTGTTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATTCTAGTTTTGAAATAATGTTCCCTAGTATACAGCAATTTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTCATATTTCATAAGGAT
AGGTCTCATTCAACCCACATGTTTATCATTTGGGATCTGCATTTTCTGATGCAAAATGATTT
ATTCTCCAGAGCAGTGAAATTTGGGTGGAATCTTATAACGGCCAAAACTAAATGCTTC
TGTGCTAAACAGAGAGTTACAAGACCCCTTTATGTGGATGGCAGCATTTTAGTCATCTTA
TGACAGAAATGTGAGAGTGGAGCTGCCACTGGGGGAGGGGGTGTGCTTGGCAGGATTTCTC
TTGGAACATCACACAAAGAAATTCCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGACTCTTGAACATGAGCTTTTAAAGTACGTTTGGCTGTTTCAGACCTTGACTTTGAG
GTGAAGGAAAAAGCTGCCAACTCCTTTTTATATGTAACCAATATATCAAGAGTCAATGTGCA
GACAGTATGCCAGTCCCAAGATCTGTCAATATGATGACTGAAGACACATTTGGGATTTATGACT
AAGGCAGGAGAAAGGCACAGCTACAGTGAAGGCCAGTTTCACTATAAAGCTTTATGAGAAATT
AGATAAGAAGGGTTTCTAATTTTTTAAATTTTTTTTATTAGATATTTTCTTCATTTACATTTT
AAATGCTATCCCAAAAGTCCCTATACTCCCCCTCACCCTGCTCCCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTTGTGATAGTGGTCAATATGATCCACCAAGCTTTACATGCTC
ACTATCTGGTCTATTGCAAGAAATGGCTGCGAGCTGATGTCAGTCAGATACAGACACCTAC
AGCCAAACAGTGGAAAGGAAGTTGGGGACTCTTATGGAAGAAAAAGGAGGAAGGGTTATGG

GCCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
AOCCTTGGGGGCTCTCAGAGTCTGAACCAACCAACCATAGAACATTCATGGGCTGTACCCAG
GCCTCTOCACTCATATGTAAACAGATATGTGGCTTGGCCTTCATCTGGGTCTGAAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCAOCTAGCTTTAGCCTT
GTAAAGACTTGAAGTTCGAGGTGTGGGTGGAGGGTATACTCAGGGAGGGCCCTCAOCTGC
TAGGAAGAGAAGAGGAGGGGGAAGACTTGGGGGAGGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAAATAAAAAATAATTAAAAAAAAAAAAAAAGAAATGGCTGCTGAGCCC
TACTCTAAAAACCATTGCATCCCCCCCCCAATCATTCACTGACTACGAATTAATAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTAACTTCTTTGTGAACAAGAGTGGTATTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTCTTTTCACTTTCCCTGTACATGCATG
TAGCTTAATGATGTTTCAGTGAATTAATAATAACAATGAAGAATAAAGACAACGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCAOCTGGAAGAAATGTCTTCAGTT
TTCTATATATGTTTACTCTATOGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACCTTATAAGAAAGTAAAAACATTCATGAAATTGTGAAGGTACTTTTC
CTTACCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTTGAA
AGAATGTAATAACCTACTAAGGCAAAACATAGAAGTTGAAATTAAGAAAGAGTAGACAGGA
GAAGTAATAAGGCAAAATAATGAATATTTGGCTTAAATAGTTCCTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAAATAATAACATTTCTACTTGAATAATTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGACCAAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

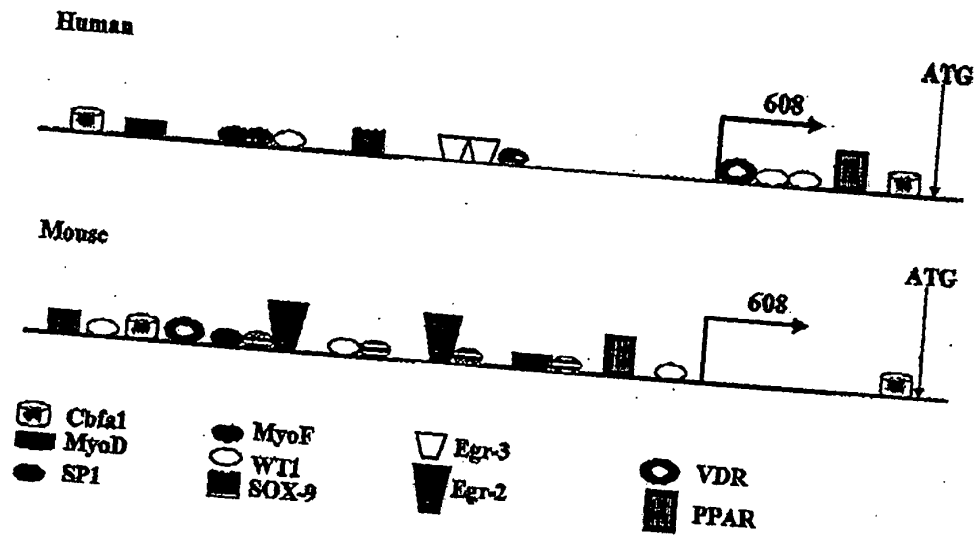


Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTCCCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAACCAACCCAGATGCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTA AAAACAACGACTTTTCATAAAATTTT
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAACTGAGTCTTAAGGATTATATTAACCTCCAGTTCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTAACCTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCAATTTAACCCCCCAA
GTAGGCATAATTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTACTGACATAATAAATAGG
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTTCGTGCCACACAGGTCTCCAGGG
TTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GCCACCTCAGCTTTGGCAGCTCCATTCTGTGGAGCCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTT
CACAAGTCTTTTTAGAAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGTAATGTGTTCATGTATGCTCATGGATGCTATTATAATGTGGGAA
TAAACTGACTTTAAAAAAGCTTATGATACTTGTACAGAGTAAATCTTCCATA
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAACCTTTAGAAA
TTTTGTTAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
CAACAAACAACAACGACAACAAACCCATAAAAAATTATCACGGCAAATGAACTAAGC
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTTCCTTAATAGTCAA
ATATGCAGAAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
GTAGATTAGAGGACACAAAGCAACAACCTACAGAGTGAAGTAATCCAGAGACTTAAT
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT
ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA
GAGGAAGAGGATAGGTGCCTCTCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
CTTGACTCTAGCTGCATATATATCAAAGATGGCCTAATCGGCCATCACTGGAAAGA
GAGGCCATTGGACTTGCAAACCTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAAATGGAAAAAAA
AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAAACCTTA
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
ACCACTACCAGCCCCAGCCCCCTTTGCCCCGCCCCAACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
TCATCTGCTCACTACCCACCCAAGCTTAATCCTTTTCCTTCCCTGCAATAGTACCTC
TCCTATTTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAAACCTCGCAGGCGAGCCTGG
GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCCATTCCCAGAAGCCGAGGCTATA
GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
CCCAAAGGCGGTGCAATTCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTACGGG
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAATGTATGATACCATTAAAATAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTTCTTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACCTCATG
TGGTATTCAATTTACGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTAAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTAAACTCATTCTATAA
CAGCTCCAGGTTGGGAGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAGAGAAGCAGAATTCTTCCTTGGCCATGCTCCCCATCACTGTGTACGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCTCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGGAAAGGTAAAGCCA
AATAAATAAAACACATATTAAATAATACAATGTTACAAAAATTGATCATATAAAGA
AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAAATATTTTTAAATG
TATATAAAAGTGTAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTCTCCAGATTTCACTAGCATTGTGTGT
CATATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTCTTTCAGTTCCTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAACAAAAACAAACCAACCAACCAAAAAACCTCTGAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTGTGCTATTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAGGTTCATTTT
GCCACAAAAAGGTGTGAATTAAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAAGTGGCCCTA
TAAGTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTTCGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCCTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTAAAAAAGCATTTCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAATACTTGCCTTACTAAATCTGACAAATTATGGTGTATTTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTTCACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTTCTTAATTAGTGATTTCATGGG
GCAGGGCCCAGTCCATTGTTCTGTTGTTACCATTTCTCAGGCACTATTAACAAAAA
AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTTCTACCTATTTGAGTTCTGTCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCAACTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTCATTTTGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCACTCCTTGCAAGTGAGATAGGTAGAGGGTGATGGAGGC
TGAGAAGAGAGGTGCAACTGTGGGTATACCCAAAAGCTGCTGATTCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCAATTGT
AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

1000
900
800
700
600
500
400
300
200
100
0

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTGAGAATCCCAGAGACT
CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC
TCAGAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
GCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAG
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAGGAAGG
AGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
GAAAGGGGGAGAAAGAAGAGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAACTACAATTATT
CCCACAAATGAAGAAAATGCTGTAAGAGTTTTACACACCCGAAGTGGGAAACTTAA
GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
CTCAACTCTTTAAGCAAACAATGCTGTTGTTTCATGAAAAGCACAATAAAGTACATAT
GTCCCATAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
CACCCACTGTTACAGAGACTTTAAACATGAACTGGAACCTATGTCTAGTGTTTGA
CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTGATGTATCCTGCATGAAATTC
TCCACCTATGTTTATTCTCTTAGCCAGGTTTTCTCTGATGGAGAAACATTGGGTTTG
AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTATTAGTATTTTCATGTTCTTT
TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA
AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAATTT
TTTGCTTCTAATTCAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAG
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCCGGCTACATG
ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
TAAAAAATAAAAAAACTATTGATGTGGAATGATCTTAATTTGGGGAAAAGACAATT
TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAATG
AGATGGTAGAAAGGAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCCAAAGTTCATGGTCAATGTGTGT
GGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATCTAA
AAACAAAACATTTTTTGACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTJAACTGCTTTTTA
CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCCCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCAATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACCTATCTTGTTTGTGTTTCTTTTTTT
ATARKACGTATTTTCTCAATTTCAATTTAGAATGATATCCCAAAAGTCCCCCATAACC
TCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Figure 51

MPKRAHWGALSVVLLLLWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
LSTIRHLYLAENMVRTLPASMLRNMPLLENLYLQGNPWTCDCEMRWFLWDKASRGIL
KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSEEEQEQ
EEDGGSQLEKFQLPQWSISLNMDEHGNMVNLVCDIKKPMDEVYKIHLNQTDPPDIDIN
ATVALDFECPMTRENYEKLWKLIAYYSEVPVKLHRELMLSKDPRVSYQYRQDADEEAL
YYTGVRQAQILAEPEWVMQPSIDIQLNRRQSTAKKVLLSYTQYSQTISTKDTRQARGRS
WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPDSKFSIL
SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDTVIGKNPGES
VTLPCNALAIPEAHLRWILPNRRIINDLANTSHVYMLPNGTSLIPKVQVSDSGYYRCVAV
NQQGADHFTVGITVTKKGSGLPSKRGRPGAKALSRVREDIVEDEGGSGMGDEENTSR
RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSLSLEVTPPPFAVSPPSASP
VQTVTSAEESSADVPLLGEEEHVLTGISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
SEKTEEITSTEGDLKGTAAPTILISEPYEPSPTLHTLDTVYEKPTHEETATEGWSAADVGSS
PEPTSSEYEPPLDAVSLAESEPMQYFDPDLETQSPDEDKMKEDTFAHLTPTPTIWNDS
STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKGKEMSQTLQGGNM
LEGDPTHSRSESEGEQESKSITLPDSTLGIMSSMSPVKKPAETTGTLLDKDTTITVTTTPR
QKVAPSSTMSTHPSRRRPNGRRRLRPNKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGTPRRKHGKRPKNHRYTPSTVSSRA
SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDMYMTTRKIYSSYPKVQETLP
VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
GTPTWNPSRTAQPGRLQTDIPVTTSGENLTDPLLKELEDVDFTSEFLSSLTVSTPFHQEE
AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSPS
TILMSLGQTTTTKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEGTOHMSGPNE
LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTPSGALPENKQFTTPRLSSTTIPLPLHM
SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSGRLPFFTNTKLSF
PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPPLLHTPQTTGSPSTN
LQNPIMVSSSTQSSISFITSSVQSSGSFHHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
TAETDVTVPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
QYMCTASNLHGLDRMVVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
AALPPVIHQEKLENISLPPGLSIHICTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTDVRYG
GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLVYKSVTDK

DAGDYLCVARNKVGDDYVVLKVDVVMKPAKIEHKEENDHKVFYGGDLKVDCVATGL
PNPEISWSLPDGS LVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVGMREEGDYTCFAEN
QVGKDEMRVRVKVVTAPATIRNKTYLAVQVPYGDVVTVACEAKGEPMPKVTWLSPTN
KVIPTSSEKYQIYQDGTLLIQKAQRSDSGNYTCLVRNSAGEDRKT VWHVNVQPPKING
NPNPITTVREIAAGGSRKLIDCKAEGIPTPRVLWAFPEGVVLPA PYYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNC S
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNA A
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKL PCTPPGAGQGRFSWTL PNGMHLEG
PQTLGRVSLLDNGTTLTVREASVFD RGTYVCRMETEYGPSVTSIPVIVIA YPPRITSEPTVI
YTRPGNTVKLNCMAMGIPKADITWELPDKSHLKAGVQARLYGNRFLHPQGS LTIQHAT
QRDAGFYKCMAKNILGSDSKTTYIHVF (SEQ ID NO: 21)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC
GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT
ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG TTGGAGCTAC
TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCGATGG AGCTTTAAGA
GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA
ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTT AGACTCTCCA
CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC
AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA
TCCGTGGACC TGCATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA
AATCCAGAGG AATTCTGAAG TGTAAAAAGG ACAAAGCTTA TGAAGGCGGT
CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC
TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAAGTGC CCCAGTGGAG
CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT
GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTTGG ACTTTGAGTG
TCCAATGACC CGAGAAAAGT ATGAAAAGCT ATGGAAATTG ATAGCATACT
ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA
CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC
AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
GTGCTACTTT CCTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
GGCTTGTAAC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA
CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT
GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
AGACGCCAG GTGCAAAGGC TCTTCCAGA GTCAGAGAAG ACATCGTGGA
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGG AGAAGAAAGC TGAAACTCTG
GAAGCATTCTG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCGT TGATTAAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTTCTGCTG TTTCTCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCCTCA GCCAGCATGG
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTAAAT GACTCCAGTA
CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
TCGAAGGAGA CCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAG AAACATACT TTTGCCTAGA
ACTGTTTCTC TGAAAACCTGA GGGCCCTTAT GATTCTTAG ATTACATGAC
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
AGGAAGAAGC TGGTTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCTA
CTGCTGCCCC GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG
TGGGGAATCC AGAAACAGAA GCAACCCAG TCAACAATGA AGGAACACAG
CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATT
CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTAATCCAAA
GTGTTTGGA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAGGCC
TCCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTACCCCG GAGACCCAG
ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
TTCTTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA
ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
GGTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
CAAATCCTAG CCTCCCATA CCAGGACGTC ACTGTCTACC TGGGAGACAC
CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTTCTT
GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GGCGTCTATA AGTGCGTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTACAT TACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCGTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCAATCCCG AGATCTCCTG GAGCCTCCA
GACGGGAGTC TGGTGAATC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCACCTCTC TTATTAGAA AGCCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATTG ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GGCGTCTATA AGTGCGTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTACAT TACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCGTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCAATCCCG AGATCTCCTG GAGCCTCCA
GACGGGAGTC TGGTGAATC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCACCTCTC TTATTAGAA AGCCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATTG ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACCTT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGTA
AGGGAAGCCA GGTTGGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCTCTG TCACTTCAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGTACCAT TTTATATGAA AAAAGAAAAA CATTCTTCC
TGGAATCAC TTTTATATA ATGTTTTATA TATATATTTT TTCCTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTAATAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTTT TCTCCAACCT CTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCTT GTATGCAAAG TTTTGTGGA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT
TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA

Figure 53

ATGAAGGTAAAGGCAGAGGAATCACCTGCTTGTGGTCTCCTTTGCTGTGATCTGCCTGGTGGCCACC
CCTGGGGGCAAGGCTGTCTCTCGCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTCAGAAAGATACCTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAATATTGAGTTTATAAACCCAGAGGTTTTTATGGGCTCAACTTTCTCCGCTGGTGCACTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCTTATACTTGTCTGATAACTTCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGGAGAAGCCAGATGTAATAAAATGCAAAAAGATAGAAGTCCCTCTAGTGCTCAG
CAGTGTCCACTTTTGATGAACCTTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCT
TTCCAGTGTGCCAAGCCAACCATTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGT
AGTTCTGCTTTCATCTCTCCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTC
ACTGAAGAAAATGACTACATCGTGCTAAATACTTCAATTTTCAACATTTTGGTGTGCAACATAGATTAC
GGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATAC TAGAAAGGAGC
CACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATT
TTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAG
CTGAACAGAATGCCACCACATTAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTA
CCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGTTA
GAACATACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACCTGCCAGGCCAAGGAGACCCACCCACAC
GTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCCTTATGTAGTGAGGATGGACGGATCCTA
ATAGACAAAAGTGGAATAATTGGAATCCAGATGGCTGATAGTTTGTACACAGGCGTATATCACTGTATA
AGCAGCAATTATGATGATGCAGATATTCTACCTATAGGATAACTGTGGTAGAACCTTTGGTCTGAAGCC
TATCAGGAAAATGGGATTATCACACAGTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACT
GGTATCCCAGATGCCTCTATTAGCTGGGTTATTCCAGGAAACATGTGCTCTATCAGTCATCAAGAGAC
AAGAAAGTTCTAAACAATGGCACATTAGAATATTACAGGTCAACCCGAAAGACCAAGGTTATTATCGC
TGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTTGTATTTTCCAAGTTTCAGTCAAGATGAAAGGACAA
AGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTT
AAGGAGCCACCAGGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCA
AGCACAAGTAAGAGGCACAACCTATCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGT
TTTAGGGAGAATTAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTTGGGCGGCACTG
TTGGAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCCACCCCAAGTG
GTCACCCAACCTCCCAACATACCTGGTGAAGAAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAA
TTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCAGCAAGGACAGTGACTGCTGACTCCAGAACAATA
TCTGATAGTCCATGACAAACATAAATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTA
CCACCTGAAGAACCACAGATTTCAAACCTGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATA
AACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTTCATCCACTGTCTTTCCACTGCTA
CTTGAGCAACTGAATTTTCAAGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCC
CCAATAACAGTAAGGACTATGATCAAGATGTCAATGTCAAAATGCTTAGTAGCACCACCAACAACTA
TTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCAC
AATCACTTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACAGCT
GCTCATTCTCAGTTTCCGATCCCTGAAATAGTACAGTTAACAATCCCGCTGTTTCAGACGCTTTGGGAGG
CAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGA
TACAGCATTTTTCAGGTCAACAACCAGAGGTTCTTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTG
CTCAATGTGACATGTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTTGTCTTTT
CCAAGTGCTGCTCCCATCACCTTCCCAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAAC
CTAGTCCAGAATCCACTATTACTTGTAGAACAACCCAGTGTAGAGAAAACAACACCCACAATAAAA
TATTTCAAGGACTGAAATTTCCCAAGTGAATCCAACTGGTGCAGTCATGACATATGTCCAACATCCATA
CCCATGGAATAAATCTCAAAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGA
GATTCAGTGATTACATCGTCACTTTTCAGGTGCTATACCAAGCCACCAATGACTATTATAGCCATTACA
AGGTTTTCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTA
AGGAATCAACATAAAGTTAGTTTACAAAAAGCAGAGCTGTGATGCTTCTTAAACATCTCCTGCTTTA
CCACAGAGACAAAGTTCCCTTTCCATTACCAACTTTCAACAAGTGATGCAAAATTCATCTAAT
ACCTTGACTACCGCTCACCACTACGACCAAAACACACAATCCTGGAAGTCTTCAACAAAAGAGGAG
CTTCCCTTCCACCCCTTAACCTATGCTTCTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATC
ATATCAACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCTACTTATGAAACC
CAAACAGAGAGATCTAGAGCACAACAATACAAGAGAACAGGAGCCTCAAAGAAGAACAGGACTGAC
CCAAACATCTCTCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCA

1000 900 800 700 600 500 400 300 200 100 0

TTCACTCATTCATTCCTCCACAGAGAAACACAACTGGGATTTCAAGCACAAATCAGTTTTCATTCAAGAAGCTCTT
AATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAAATTGCTTCT
GAAACAACTTTGTCCAGCAAATCACACCAGAGTACCACAAC TAGGAAAGCATCATTAGACACTCCCAT
CCACCATTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCCATCTCCCCCTCCCTTTACTCAGAGAGCA
GTTACTGACACACGTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCC
TCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTA
CATTCTACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCTCCCTGG
GCAGAATACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGC
ATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACCTCTTGTTTCAGATTGGGATGGACAGAAGAACACA
AAGAAGAGTGACTTTGTATAAGAAACAGTTTCAAGAAGCAACAACCTCCAAACTCCTTCCCTTTGACTCT
TTGCTTAGGTATATATTTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTACTATTCCAGCT
AACTCAGATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCTGCCCACCATTCTATTGGACCGAGGTT
TCAGGACTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCGTGTCC
ATCCAGAGGGTGGAAATT CAGGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGAC
CACCTTCATGTACCTTGTCTGTGGTTTCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATC
ACAGTTCATTCCGGAAGCACTGTGGAAC TGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAAAACCAACAGTTGTCTCAGAATCATCCAGGGAAGTAGGCAGGCTGTGGTGACGGTT
GACGGAACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAAC
CCAGGTGGCCAGGATTCAGTGTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAA
AGGAGGCAAGTCATTGTAGGCACCTTGGGGTGAAAGTTTAAAACTGCCCTGTACTGCAAAAGGAACCTCT
CAGCCCAGCGTTTACTGGGTCTCTCTGATGGCACTGAAAGTGAACCATTACAGTTTACCAATTCCAAG
TTGTTCTTATTTTCAAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACCTTATGAA
TGCATTGCTACCAGTTCCTACTGGTTCCGAGGCCAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACC
AGCCCAGGATAGAAGCTGCATCCCGAAAAGGACTGAAGTGAATTTGGGGACAAATTACTACTGAAC
TGCTCAGCCACTGGGGAGCCCAAACCCCAATAATGTGGAGGTTACCATTCCAAGGCTGTGGTTCGACCAG
TGGAGCTGGATCCACGTCTACCCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGT
GTCTACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTG
AAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTA
GATTGCAAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTGGAGTTTGCCTGATGGAACCATGATCAAC
AATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATAACCTTTTCAACAATGGAACTTTA
TACTTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTACTTGTCTATGCCCAGAACACCCTAGGGAAA
GATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAGTAACAAAACCAAC
AAGAGAATCAAAGCTGGAGACACAGCTGTCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAAATA
TTTTGGTTGCTGCCTTCCAATGACATGATTTCCCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCGAAATCCCAGT
GGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAAATCAATGGTCTGTATACA
AACAGACTGTTTATTAAGGCCACAGCTGTGAGACATCTCAAAAACACTTTGACTGCAGAGCTGAAGGG
ACACCATCTCTGGAAGTCATGTGGATCATGCCAGACATAATTTTCTCACAGCCCCATAC TATGGAAGC
AGAATCACAGTCCATAAAAATGGAACCTTGAAATTAGGAATGTGAGGCTTTT CAGATT CAGCCGACTTT
ATCTGTGTGGCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAATGCTG
AGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGAAAGTCCACAGCATTG
AATTGCTCTGTTGATGGTAACCCACCACCTGAAATAACTCTGGATTTTACCAATGGCACACGATTTTCC
AATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTCTTTTATCATTTCTAAAACAACCTCGG
GAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTA
GAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGAATCT
CTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAAATATCAAATGGACTATGCCAAGTGGTTAT
GTAGTAGACAGGCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAA
GCAACAGCTTATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTGGTGCATACACTGATTACT
GTTCCAGTAATGATTGTAGCCTACCCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGG
ACAGGGGCAGCCTTT CAGCTCCACTGTGTGGCTTTGGGAGTTC CCAAGCCAGAAATCACATGGGAGATG
CTTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGGACACATGGAAGTGAGCAGCTTCATCTTACAA
GGTACCTTAGTCATT CAGAATCCCCAACCTCCGATTTCTGGGATATACAAATGCACAGCAAAGAACCCA
CTTGGTAGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 54

ORIGIN

1	MKVKGRCITC	LLVSFAVICL	VATPGGKACP	RRCACYMPT	VHCTFRYLTS
51	IPDSIPPVNE	RINLGYNLSV	RLMETDFSG	TKLELLMLHS	NGIHTIPDKT
101	FSDLQALQVL	KMSYNKVRKL	QKDTFYGLRS	LTRLHMDHNN	IEFINPEVFY
151	GLNFLRLVHL	EGNQLTKLHP	DTFVSLSYLQ	IFKISFIKFL	YLSDNFLTSL
201	PQEMVSYPMD	LDSLYLHGNP	WTCDCHLKWL	SDWIEKPDV	IKCKKDRSPS
251	SAQQCPLCMN	PRTSKGKPLA	MVSAAAFQCA	KPTIDSSLKS	KSLTILEDSS
301	SAFISPOGFM	APFGSLTLNM	TDQSGNEANM	VCSIQKPSRT	SPIAFTEEND
351	YIVLNTSFST	FLVCNIDYGH	IQPVWQILAL	YSDSPLILER	SHLLSETPQL
401	YYKYQVAPK	PEDIFTNIEA	DLRADPSWLM	QDQISLQLNR	TATTFSTLQI
451	QYSSDAQITL	PRAEMRPVKH	KWTMISRDNM	TKLEHTVLVG	GTVGLNCPGQ
501	GDPTPHVDWL	LADGSKVRAP	YVSEDGRILI	DKSGKLELQM	ADSFDTGVYH
551	CISSNYDDAD	ILTYRITVVE	PLVEAYQENG	IHHTVFIGET	LDLPCHSTGI
601	PDASISWVIP	GNNVLYQSSR	DKKVLNNGTL	RILQVTPKDQ	GYRCVAANP
651	SGVDFLIFQV	SVKMKQGRPL	EHDGETEGSG	LDESNPJIAHL	KEPPGAQLRT
701	SALMEAEVVK	HTSSTSKRHN	YRELTQLRRG	DSTHRRFREN	RRHFPPSARR
751	IDPQHWALL	EKAKKNAMPD	KRENTTVSPP	PVVTQLPNIP	GEEDSSGML
801	ALHEEFMVPA	TKALNLPART	VTADSRTISD	SPMTNINYG	EFSPVNSQI
851	LPPEPTDFK	LSTAIKTAM	SKNINPTMSS	QIQGTTNQH	STVFPLLLGA
901	TEFQDSQMG	RGREHFQSRP	PITVRTMIKD	VNVKMLSST	NKLLLESVNT
951	TNSHQTSVRE	VSEPRHNHFY	SHTTQILSTS	TFPSDPHTAA	HSQFPPIRNS
1001	TVNIPLFRFR	GRQRKIGGRG	RIISPYRTPV	LRRHYSIFR	STTRGSSEKS
1051	TTAFSATVLN	VTCLSCLPRE	RLTTATAALS	FPSAAPITFP	KADIARVPSE
1101	ESTTLVQNPL	LLLENKPSVE	KTTPTIKYFR	TEISQVTPG	AVMTYAPTSI
1151	PMEKTHKVNA	SYPRVSSSTNE	AKRDSVITSS	LSGAITKPPM	TIIAITRFSR
1201	RKIPWQQNFV	NNHNPKGRLR	NQHKVSLQKS	TAVMLPKTSP	ALPQRQSSPF
1251	HFTTLSTSV	QIPSNLTLLTA	HHTTTKTHNP	GSLPTKKELP	FPPLNMLPS
1301	IISKDSSTKS	IISTQTAIPA	TTPTFPASVI	TYETQTERS	AQTIQREQEP
1351	QKKNRDPNI	SPDQSSGFTT	PTAMTPPALA	FTHSPPENTT	GISSTISFHS
1401	RTLNLTDVIE	ELAQASTQTL	KSTIASSETL	SSKSHQSTTT	RKASLDTPIP
1451	PFLSSSATLM	PVPISPPFTQ	RAVTDTRGDS	HFRLMTNTVV	KLHESSRHNL
1501	QMPSSQLEPL	TSSTSLLLHS	TPMPALTTVK	SQNSKLTSP	WAEYQFWHKP
1551	YSDIAEKGGK	PEVSMLATG	LSEATTLVSD	WDGQKNTKKS	DFDKKPVQEA
1601	TTSKLLPFDS	LSRYIFEKPR	IVGGKAASFT	IPANSDAFLP	CEAVGNPLPT
1651	IHWTRVSGLD	LSRGNQNSRV	QVLPNGTSLI	QRVEIQDRGQ	YLCASNLFG
1701	TDHLHVTLSV	VSYPRIER	RTKEITVHSG	STVELKCRAE	GRPSPTVTWI
1751	LANQTVVSES	SQGSRQAVVT	VDGTLVLHNL	SIYDRGFYKC	VASNPGGQDS
1801	LLVKIQVIAA	PPVILEQRRQ	VIVGTWGESL	KLPCTAKGTP	QPSVYWVLS
1851	GTEVKPLQFT	NSKFLFSSNG	TLYIRNLASS	DRGTYECIAT	SSTGSERRVV
1901	MLTMEERVTS	PRIEASQKR	TEVNFGDKLL	LNCSATGEPK	PQIMWRLPSK
1951	AVVDQWSWIH	VYPNGSLFIG	SVTEKDSGVY	LCVARNKMGD	DLILMHVSLR
2001	LKPAKIDHKQ	YFRKQVLHGK	DFQVDCASG	SPVPEISWSL	PDGTMNNAM
2051	QADDSGHRTR	RYTLFNNGTL	YFNKVGVAEE	GDYTCYAQNT	LGKDEMKVHL
2101	TVITAAPRIR	QSNKTNKRIK	AGDTAVLDCE	VTGDPKPKIF	WLLPSNDMIS
2151	FSIDRYTFHA	NGSLTINKVK	LLDSGEYVCV	ARNPSGDDTK	MYKLDVVS
2201	PLINGLYTNR	TVIKATAVRH	SKKHFDCAE	GTPSPEVMWI	MPDNIFLTAP
2251	YYGSRITVHK	NGTLEIRNVR	LSDSADFCV	ARNEGGESVL	VVQLEVLEML
2301	RRPTFRNPFN	EKIVAQLGKS	TALNCSVDGN	PPPEIIWILP	NGTRFSNGPQ
2351	SYQYLIASNG	SFIISKTTRE	DAGKYRCAAR	NKVGYIEKLV	ILEIGQKQVI
2401	LTYAPGTVKG	ISGESLSLHC	VSDGIPKPN	KWTMPSGYVV	DRPQINGKYI
2451	LHDNGTLVIK	EATAYDRGNY	ICKAQNSVGH	TLITVPVMIV	AYPPRITNRP
2501	PRSIIVTRTGA	AFQLHCVALG	VPKPEITWEM	PDHSLSTAS	KERTHGSEQL
2551	HLQGTLLVIQN	PQTSDSGIYK	CTAKNPLGSD	YAATYIQVI*	

Figure 55 **OCP rat amino acid sequence**

>608-663Nterm Rat Protein (663 aa)

MQVRGREVSGLLISLTAVCLVVTGPSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVLSYLQ
IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNG
SASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
NLVCSVDYNHIQPVWQLLALYSDSLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
DVRADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNN
PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQM
ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
SVQ

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663

Figure 56

pCM-H-608-663-N-term

